

US010097919B2

(12) **United States Patent**  
**Jarvis et al.**

(10) **Patent No.:** **US 10,097,919 B2**  
(45) **Date of Patent:** **Oct. 9, 2018**

(54) **MUSIC SERVICE SELECTION**

(71) Applicant: **SONOS, INC.**, Santa Barbara, CA (US)

(72) Inventors: **Simon Jarvis**, Cambridge, MA (US);  
**Mark Plagge**, Santa Barbara, CA (US);  
**Christopher Butts**, Evanston, IL (US)

(73) Assignee: **SONOS, INC.**, Santa Barbara, CA (US)

(\*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **15/098,718**

(22) Filed: **Apr. 14, 2016**

(65) **Prior Publication Data**

US 2017/0242649 A1 Aug. 24, 2017

**Related U.S. Application Data**

(60) Provisional application No. 62/298,410, filed on Feb. 22, 2016, provisional application No. 62/298,418, (Continued)

(51) **Int. Cl.**

**G10L 21/02** (2013.01)  
**H04R 3/00** (2006.01)  
**G06F 3/16** (2006.01)  
**H04R 29/00** (2006.01)  
**H04W 8/00** (2009.01)  
**H04W 8/24** (2009.01)  
**H04R 27/00** (2006.01)  
**H04L 12/28** (2006.01)  
**G10L 15/14** (2006.01)

(Continued)

(52) **U.S. Cl.**

CPC ..... **H04R 3/00** (2013.01); **G06F 3/162** (2013.01); **G06F 3/165** (2013.01); **G06F 3/167** (2013.01); **G10L 15/14** (2013.01); **G10L**

**15/22** (2013.01); **H04L 12/2803** (2013.01); **H04R 3/12** (2013.01); **H04R 27/00** (2013.01); **H04R 29/007** (2013.01); **H04S 7/301** (2013.01); **H04S 7/303** (2013.01); **H04W 8/005** (2013.01); **H04W 8/24** (2013.01); **G10L 21/02** (2013.01); **G10L 2015/223** (2013.01); **H04L 2012/2849** (2013.01); **H04R 2227/005** (2013.01); **H04R 2420/07** (2013.01)

(58) **Field of Classification Search**

None  
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,741,038 A 4/1988 Elko et al.  
4,941,187 A 7/1990 Slater  
(Continued)

FOREIGN PATENT DOCUMENTS

AU 2017100486 A4 6/2017  
AU 2017100581 A4 6/2017  
(Continued)

OTHER PUBLICATIONS

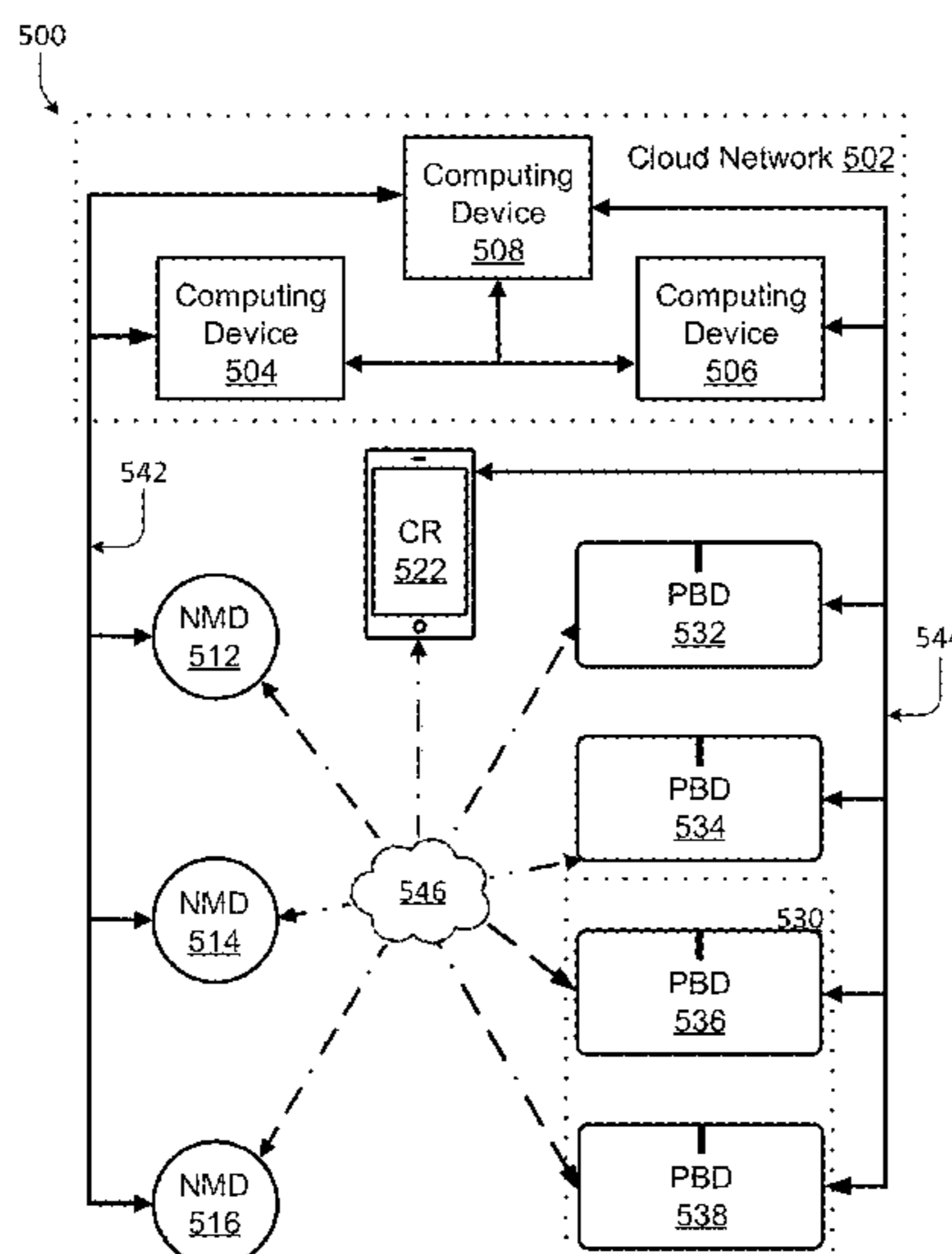
US 6,489,633, 10/2002, Wachter (withdrawn)  
(Continued)

*Primary Examiner* — Michael N Opsasnick

(57) **ABSTRACT**

Methods and apparatus for identifying a music service based on a user command. A content type is identified from a received user command and a music service is selected that supports the content type. A selected music service can then transmit audio content associated with the content type for playback.

**11 Claims, 7 Drawing Sheets**



**Related U.S. Application Data**

filed on Feb. 22, 2016, provisional application No. 62/298,433, filed on Feb. 22, 2016, provisional application No. 62/298,439, filed on Feb. 22, 2016, provisional application No. 62/298,425, filed on Feb. 22, 2016, provisional application No. 62/298,350, filed on Feb. 22, 2016, provisional application No. 62/298,388, filed on Feb. 22, 2016, provisional application No. 62/298,393, filed on Feb. 22, 2016, provisional application No. 62/312,350, filed on Mar. 23, 2016.

**(51) Int. Cl.****G10L 15/22** (2006.01)**H04R 3/12** (2006.01)**H04S 7/00** (2006.01)**(56)****References Cited****U.S. PATENT DOCUMENTS**

5,440,644 A 8/1995 Farinelli et al.  
5,740,260 A 4/1998 Odom  
5,761,320 A 6/1998 Farinelli et al.  
5,923,902 A 7/1999 Inagaki  
6,032,202 A 2/2000 Lea et al.  
6,256,554 B1 7/2001 DiLorenzo  
6,301,603 B1 10/2001 Maher et al.  
6,311,157 B1 10/2001 Strong  
6,404,811 B1 6/2002 Cvetko et al.  
6,522,886 B1 2/2003 Youngs et al.  
6,594,347 B1 7/2003 Calder et al.  
6,611,537 B1 8/2003 Edens et al.  
6,631,410 B1 10/2003 Kowalski et al.  
6,757,517 B2 6/2004 Chang  
6,778,869 B2 8/2004 Champion  
7,130,608 B2 10/2006 Hollstrom et al.  
7,130,616 B2 10/2006 Janik  
7,143,939 B2 12/2006 Henzerling  
7,236,773 B2 6/2007 Thomas  
7,295,548 B2 11/2007 Blank et al.  
7,391,791 B2 6/2008 Balassanian et al.  
7,483,538 B2 1/2009 McCarty et al.  
7,571,014 B1 8/2009 Lambourne et al.  
7,630,501 B2 12/2009 Blank et al.  
7,643,894 B2 1/2010 Braithwaite et al.  
7,657,910 B1 2/2010 McAulay et al.  
7,661,107 B1 2/2010 Van Dyke et al.  
7,702,508 B2 4/2010 Bennett  
7,792,311 B1 9/2010 Holmgren et al.  
7,853,341 B2 12/2010 McCarty et al.  
7,987,294 B2 7/2011 Bryce et al.  
8,014,423 B2 9/2011 Thaler et al.  
8,041,565 B1 10/2011 Bhardwaj et al.  
8,045,952 B2 10/2011 Qureshey et al.  
8,073,125 B2 12/2011 Zhang et al.  
8,103,009 B2 1/2012 McCarty et al.  
8,234,395 B2 7/2012 Millington et al.  
8,239,206 B1\* 8/2012 LeBeau ..... H04M 1/271  
704/251  
8,255,224 B2 8/2012 Singleton et al.  
8,290,603 B1 10/2012 Lambourne  
8,364,481 B2 1/2013 Strope et al.  
8,386,261 B2 2/2013 Mellott et al.  
8,423,893 B2 4/2013 Ramsay et al.  
8,428,758 B2 4/2013 Naik et al.  
8,453,058 B1 5/2013 Coccaro et al.  
8,483,853 B1 7/2013 Lambourne  
8,484,025 B1 7/2013 Moreno et al.  
8,831,761 B2 9/2014 Kemp et al.  
8,831,957 B2 9/2014 Taubman et al.  
8,938,394 B1 1/2015 Faaborg et al.  
8,942,252 B2 1/2015 Balassanian et al.  
8,983,383 B1 3/2015 Haskin  
8,983,844 B1 3/2015 Thomas et al.

9,042,556 B2 5/2015 Kallai et al.  
9,094,539 B1 7/2015 Noble  
9,215,545 B2 12/2015 Dublin et al.  
9,251,793 B2 2/2016 Lebeau et al.  
9,253,572 B2 2/2016 Beddingfield, Sr. et al.  
9,262,612 B2 2/2016 Cheyer  
9,288,597 B2 3/2016 Carlsson et al.  
9,300,266 B2 3/2016 Grokop  
9,318,107 B1 4/2016 Sharifi  
9,319,816 B1 4/2016 Narayanan  
9,412,392 B2 8/2016 Lindahl et al.  
9,426,567 B2 8/2016 Lee et al.  
9,431,021 B1\* 8/2016 Scalise ..... G10L 21/00  
9,443,527 B1 9/2016 Watanabe et al.  
9,472,201 B1 10/2016 Sleator  
9,472,203 B1 10/2016 Ayrapetian et al.  
9,484,030 B1 11/2016 Meaney et al.  
9,489,948 B1 11/2016 Chu et al.  
9,494,683 B1 11/2016 Sadek  
9,509,269 B1 11/2016 Rosenberg  
9,514,752 B2 12/2016 Sharifi  
9,536,541 B2 1/2017 Chen et al.  
9,548,066 B2 1/2017 Jain et al.  
9,552,816 B2 1/2017 Vanlund et al.  
9,560,441 B1 1/2017 McDonough, Jr. et al.  
9,601,116 B2 3/2017 Casado et al.  
9,615,170 B2 4/2017 Kirsch et al.  
9,615,171 B1 4/2017 O'Neill et al.  
9,632,748 B2 4/2017 Faaborg et al.  
9,633,186 B2 4/2017 Ingrassia, Jr. et al.  
9,633,368 B2 4/2017 Greenzeiger et al.  
9,633,660 B2 4/2017 Haughay et al.  
9,633,671 B2 4/2017 Giacobello et al.  
9,633,674 B2 4/2017 Sinha et al.  
9,640,179 B1 5/2017 Hart et al.  
9,640,183 B2 5/2017 Jung et al.  
9,641,919 B1 5/2017 Poole et al.  
9,646,614 B2 5/2017 Bellegarda et al.  
9,653,060 B1 5/2017 Hilmes et al.  
9,653,075 B1 5/2017 Chen et al.  
9,659,555 B1 5/2017 Hilmes et al.  
9,672,821 B2 6/2017 Krishnaswamy et al.  
9,685,171 B1 6/2017 Yang  
9,691,378 B1 6/2017 Meyers et al.  
9,691,379 B1 6/2017 Mathias et al.  
9,697,826 B2 7/2017 Sainath et al.  
9,697,828 B1 7/2017 Prasad et al.  
9,698,999 B2 7/2017 Mutagi et al.  
9,704,478 B1 7/2017 Vitaladevuni et al.  
9,721,568 B1 8/2017 Polansky et al.  
9,721,570 B1 8/2017 Beal et al.  
9,728,188 B1 8/2017 Rosen et al.  
9,734,822 B1 8/2017 Sundaram et al.  
9,747,011 B2 8/2017 Lewis et al.  
9,747,899 B2 8/2017 Pogue et al.  
9,747,920 B2 8/2017 Ayrapetian et al.  
9,747,926 B2 8/2017 Sharifi et al.  
9,754,605 B1 9/2017 Chhetri  
9,762,967 B2 9/2017 Clarke et al.  
9,811,314 B2 11/2017 Plagge et al.  
9,813,810 B1 11/2017 Nongpiur  
9,820,036 B1 11/2017 Tritschler et al.  
9,820,039 B2 11/2017 Lang  
9,826,306 B2 11/2017 Lang  
9,881,616 B2 1/2018 Beckley et al.  
2001/0042107 A1 11/2001 Palm  
2002/0022453 A1 2/2002 Balog et al.  
2002/0026442 A1 2/2002 Lipscomb et al.  
2002/0034280 A1 3/2002 Infosino  
2002/0072816 A1 6/2002 Shdema et al.  
2002/0124097 A1 9/2002 Isely et al.  
2003/0040908 A1 2/2003 Yang et al.  
2003/0072462 A1 4/2003 Hlibowicki  
2003/0157951 A1 8/2003 Hasty  
2004/0024478 A1 2/2004 Hans et al.  
2005/0164664 A1 7/2005 Difonzo et al.  
2005/0195988 A1 9/2005 Tashev et al.  
2005/0268234 A1 12/2005 Rossi et al.  
2005/0283330 A1 12/2005 Laraia et al.

(56)

References Cited

U.S. PATENT DOCUMENTS

2006/0147058 A1 7/2006 Wang  
 2006/0190968 A1 8/2006 Jung et al.  
 2006/0247913 A1 11/2006 Huerta et al.  
 2006/0262943 A1 11/2006 Oxford  
 2007/0018844 A1 1/2007 Sutardja  
 2007/0019815 A1 1/2007 Asada et al.  
 2007/0033043 A1 2/2007 Hyakumoto  
 2007/0071255 A1 3/2007 Schobben  
 2007/0076131 A1 4/2007 Li et al.  
 2007/0140058 A1 6/2007 McIntosh et al.  
 2007/0142944 A1 6/2007 Goldberg et al.  
 2008/0090537 A1 4/2008 Sutardja  
 2008/0248797 A1 10/2008 Freeman et al.  
 2008/0301729 A1 12/2008 Broos et al.  
 2009/0003620 A1 1/2009 McKillop et al.  
 2009/0005893 A1 1/2009 Sugii et al.  
 2009/0018828 A1 1/2009 Nakadai et al.  
 2009/0076821 A1\* 3/2009 Brenner ..... G06F 17/30053  
 704/260  
 2009/0197524 A1 8/2009 Haff et al.  
 2009/0228919 A1 9/2009 Zott et al.  
 2009/0238377 A1 9/2009 Ramakrishnan et al.  
 2009/0326949 A1\* 12/2009 Douthitt ..... G06F 17/30038  
 704/260  
 2010/0014690 A1 1/2010 Wolff et al.  
 2010/0023638 A1 1/2010 Bowman  
 2010/0075723 A1 3/2010 Min et al.  
 2010/0179874 A1 7/2010 Higgins et al.  
 2010/0211199 A1 8/2010 Naik et al.  
 2011/0033059 A1 2/2011 Bhaskar et al.  
 2011/0091055 A1 4/2011 Leblanc  
 2011/0145581 A1 6/2011 Malhotra et al.  
 2011/0267985 A1 11/2011 Wilkinson et al.  
 2011/0276333 A1 11/2011 Wang et al.  
 2011/0280422 A1 11/2011 Neumeyer et al.  
 2011/0299706 A1 12/2011 Sakai  
 2012/0131125 A1 5/2012 Seidel et al.  
 2012/0148075 A1 6/2012 Goh et al.  
 2012/0163603 A1 6/2012 Abe et al.  
 2012/0177215 A1 7/2012 Bose et al.  
 2012/0297284 A1 11/2012 Matthews, III et al.  
 2013/0006453 A1 1/2013 Wang et al.  
 2013/0024018 A1 1/2013 Chang et al.  
 2013/0058492 A1 3/2013 Silzle et al.  
 2013/0066453 A1 3/2013 Seefeldt  
 2013/0148821 A1 6/2013 Sorensen  
 2013/0179173 A1 7/2013 Lee et al.  
 2013/0183944 A1 7/2013 Mozer et al.  
 2013/0191122 A1\* 7/2013 Mason ..... G06F 17/30017  
 704/231  
 2013/0216056 A1 8/2013 Thyssen  
 2013/0317635 A1 11/2013 Bates et al.  
 2013/0329896 A1 12/2013 Krishnaswamy et al.  
 2013/0331970 A1 12/2013 Beckhardt et al.  
 2013/0332165 A1 12/2013 Beckley et al.  
 2013/0343567 A1 12/2013 Triplett et al.  
 2014/0003611 A1 1/2014 Mohammad et al.  
 2014/0003635 A1 1/2014 Mohammad et al.  
 2014/0006026 A1 1/2014 Lamb et al.  
 2014/0064501 A1 3/2014 Olsen et al.  
 2014/0075306 A1\* 3/2014 Rega ..... G06F 17/30749  
 715/716  
 2014/0094151 A1 4/2014 Klappert et al.  
 2014/0100854 A1 4/2014 Chen et al.  
 2014/0167931 A1 6/2014 Lee et al.  
 2014/0195252 A1 7/2014 Gruber et al.  
 2014/0244013 A1 8/2014 Reilly  
 2014/0258292 A1\* 9/2014 Thramann ..... G06F 17/30053  
 707/736  
 2014/0270282 A1 9/2014 Tammi et al.  
 2014/0274185 A1 9/2014 Luna et al.  
 2014/0277650 A1 9/2014 Zurek et al.  
 2014/0363022 A1 12/2014 Dizon et al.  
 2015/0010169 A1 1/2015 Popova et al.

2015/0016642 A1 1/2015 Walsh et al.  
 2015/0063580 A1 3/2015 Huang et al.  
 2015/0086034 A1 3/2015 Lombardi et al.  
 2015/0104037 A1 4/2015 Lee et al.  
 2015/0154976 A1 6/2015 Mutagi  
 2015/0180432 A1 6/2015 Gao et al.  
 2015/0189438 A1 7/2015 Hampiholi et al.  
 2015/0200454 A1 7/2015 Heusdens et al.  
 2015/0222987 A1 8/2015 Angel, Jr. et al.  
 2015/0228274 A1 8/2015 Leppänen et al.  
 2015/0253292 A1 9/2015 Larkin et al.  
 2015/0253960 A1 9/2015 Lin et al.  
 2015/0271593 A1 9/2015 Sun et al.  
 2015/0280676 A1 10/2015 Holman et al.  
 2015/0296299 A1 10/2015 Klippel et al.  
 2015/0302856 A1 10/2015 Kim et al.  
 2015/0341406 A1\* 11/2015 Rockefeller ..... H04L 65/403  
 709/219  
 2015/0363061 A1\* 12/2015 de Nigris, III ..... H04N 21/8113  
 715/717  
 2015/0363401 A1\* 12/2015 Chen ..... G06F 17/3053  
 707/723  
 2015/0371657 A1 12/2015 Gao et al.  
 2015/0380010 A1 12/2015 Srinivasan  
 2016/0007116 A1 1/2016 Holman  
 2016/0021458 A1 1/2016 Johnson et al.  
 2016/0029142 A1 1/2016 Isaac  
 2016/0036962 A1 2/2016 Rand et al.  
 2016/0042748 A1 2/2016 Jain et al.  
 2016/0057522 A1 2/2016 Choisel et al.  
 2016/0077710 A1 3/2016 Lewis et al.  
 2016/0088392 A1 3/2016 Huttunen et al.  
 2016/0093304 A1 3/2016 Kim et al.  
 2016/0098393 A1\* 4/2016 Hebert ..... G06F 17/28  
 704/9  
 2016/0157035 A1 6/2016 Russell et al.  
 2016/0173578 A1\* 6/2016 Sharma ..... H04L 67/1002  
 709/203  
 2016/0212538 A1 7/2016 Fullam et al.  
 2016/0225385 A1 8/2016 Hammarqvist  
 2016/0232451 A1 8/2016 Scherzer  
 2016/0234204 A1 8/2016 Rishi et al.  
 2016/0239255 A1 8/2016 Chavez et al.  
 2016/0260431 A1 9/2016 Newendorp et al.  
 2016/0314782 A1 10/2016 Klimanis  
 2016/0352915 A1 12/2016 Gautama  
 2016/0353218 A1 12/2016 Starobin et al.  
 2017/0003931 A1 1/2017 Dvortsov et al.  
 2017/0026769 A1 1/2017 Patel  
 2017/0060526 A1 3/2017 Barton et al.  
 2017/0070478 A1 3/2017 Park et al.  
 2017/0076720 A1 3/2017 Gopalan et al.  
 2017/0078824 A1 3/2017 Heo  
 2017/0084292 A1 3/2017 Yoo  
 2017/0090864 A1 3/2017 Jorgovanovic  
 2017/0092278 A1 3/2017 Evermann et al.  
 2017/0092297 A1 3/2017 Sainath et al.  
 2017/0103755 A1 4/2017 Jeon et al.  
 2017/0125037 A1 5/2017 Shin  
 2017/0177585 A1 6/2017 Rodger et al.  
 2017/0178662 A1 6/2017 Ayrapetian et al.  
 2017/0193999 A1 7/2017 Aleksic et al.  
 2017/0206896 A1 7/2017 Ko et al.  
 2017/0214996 A1 7/2017 Yeo  
 2017/0236512 A1 8/2017 Williams et al.  
 2017/0242651 A1 8/2017 Lang et al.  
 2017/0242653 A1 8/2017 Lang et al.  
 2017/0270919 A1 9/2017 Parthasarathi et al.  
 2017/0353789 A1 12/2017 Kim et al.  
 2018/0033428 A1 2/2018 Kim et al.

FOREIGN PATENT DOCUMENTS

EP 1349146 A1 10/2003  
 EP 1389853 A1 2/2004  
 EP 2351021 B1 9/2017  
 JP 2001236093 A 8/2001  
 JP 2004347943 A 12/2004

(56)

**References Cited**

## FOREIGN PATENT DOCUMENTS

JP	2004354721	A	12/2004
JP	2005284492	A	10/2005
JP	2008079256	A	4/2008
JP	2008158868	A	7/2008
JP	2010141748	A	6/2010
JP	2013037148	A	2/2013
JP	2014071138	A	4/2014
JP	2014137590	A	7/2014
KR	20100111071	A	10/2010
WO	200153994		7/2001
WO	2003093950	A2	11/2003
WO	2015037396	A1	3/2015
WO	2015178950	A1	11/2015
WO	2016033364	A1	3/2016
WO	2017039632	A1	3/2017

## OTHER PUBLICATIONS

US 9,299,346, 03/2016, Hart et al. (withdrawn)  
 "AudioTron Quick Start Guide, Version 1.0", Voyetra Turtle Beach, Inc., Mar. 2001, 24 pages.  
 "AudioTron Reference Manual, Version 3.0", Voyetra Turtle Beach, Inc., May 2002, 70 pages.  
 "AudioTron Setup Guide, Version 3.0", Voyetra Turtle Beach, Inc., May 2002, 38 pages.  
 Bluetooth. "Specification of the Bluetooth System: The ad hoc SCATTERNET for affordable and highly functional wireless connectivity," Core, Version 1.0 A, Jul. 26, 1999, 1068 pages.  
 Bluetooth. "Specification of the Bluetooth System: Wireless connections made easy," Core, Version 1.0 B, Dec. 1, 1999, 1076 pages.  
 Dell, Inc. "Dell Digital Audio Receiver: Reference Guide," Jun. 2000, 70 pages.  
 Dell, Inc. "Start Here," Jun. 2000, 2 pages.  
 Jo et al., "Synchronized One-to-many Media Streaming with Adaptive Playout Control," Proceedings of SPIE, 2002, pp. 71-82, vol. 4861.  
 Jones, Stephen, "Dell Digital Audio Receiver: Digital upgrade for your analog stereo" Analog Stereo Jun. 24, 2000 retrieved Jun. 18, 2014, 2 pages.  
 Louderback, Jim, "Affordable Audio Receiver Furnishes Homes With MP3," TechTV Vault. Jun. 28, 2000 retrieved Jul. 10, 2014, 2 pages.  
 Palm, Inc., "Handbook for the Palm VII Handheld," May 2000, 311 pages.  
 Presentations at WinHEC 2000, May 2000, 138 pages.  
 UPnP; "Universal Plug and Play Device Architecture," Jun. 8, 2000; version 1.0; Microsoft Corporation; pp. 1-54.  
 Non-Final Office dated Jan. 13, 2017, issued in connection with U.S. Appl. No. 15/098,805, filed Apr. 14, 2016, 11 pages.  
 United States Patent and Trademark Office, U.S. Appl. No. 60/490,768, filed Jul. 28, 2003, entitled "Method for synchronizing audio playback between multiple networked devices," 13 pages.  
 United States Patent and Trademark Office, U.S. Appl. No. 60/825,407, filed Sep. 12, 2006, entitled "Controlling and manipulating groupings in a multi-zone music or media system," 82 pages.  
 Yamaha DME 64 Owner's Manual; copyright 2004, 80 pages.  
 Yamaha DME Designer 3.5 setup manual guide; copyright 2004, 16 pages.  
 Yamaha DME Designer 3.5 User Manual; Copyright 2004, 507 pages.  
 "Denon 2003-2004 Product Catalog," Denon, 2003-2004, 44 pages.  
 Corrected Notice of Allowability dated Mar. 8, 2017, issued in connection with U.S. Appl. No. 15/229,855, filed Aug. 5, 2016, 6 pages.  
 Non-Final Office Action dated Feb. 7, 2017, issued in connection with U.S. Appl. No. 15/131,244, filed Apr. 18, 2016, 12 pages.  
 Non-Final Office Action dated Feb. 8, 2017, issued in connection with U.S. Appl. No. 15/098,892, filed Apr. 14, 2016, 17 pages.  
 Non-Final Office Action dated Mar. 9, 2017, issued in connection with U.S. Appl. No. 15/098,760, filed Apr. 14, 2016, 13 pages.

Non-Final Office Action dated Apr. 19, 2017, issued in connection with U.S. Appl. No. 15/131,776, filed Apr. 18, 2016, 12 pages.  
 Non-Final Office Action dated Jan. 26, 2017, issued in connection with U.S. Appl. No. 15/098,867, filed Apr. 14, 2016, 16 pages.  
 Notice of Allowance dated Feb. 14, 2017, issued in connection with U.S. Appl. No. 15/229,855, filed Aug. 5, 2016, 11 pages.  
 Final Office Action dated Aug. 11, 2017, issued in connection with U.S. Appl. No. 15/131,776, filed Apr. 18, 2016, 7 pages.  
 International Searching Authority, International Search Report and Written Opinion dated May 23, 2017, issued in connection with International Application No. PCT/US2017/018739, Filed on Feb. 21, 2017, 10 pages.  
 International Searching Authority, International Search Report and Written Opinion dated May 30, 2017, issued in connection with International Application No. PCT/US2017/018728, Filed on Feb. 21, 2017, 11 pages.  
 Non-Final Office Action dated Jun. 1, 2017, issued in connection with U.S. Appl. No. 15/223,218, filed Jul. 29, 2016, 7 pages.  
 Non-Final Office Action dated Jul. 25, 2017, issued in connection with U.S. Appl. No. 15/273,679, filed Jul. 22, 2016, 11 pages.  
 Non-Final Office Action dated Jun. 30, 2017, issued in connection with U.S. Appl. No. 15/277,810, filed Sep. 27, 2016, 13 pages.  
 Notice of Allowance dated Jul. 12, 2017, issued in connection with U.S. Appl. No. 15/098,805, filed Apr. 14, 2016, 8 pages.  
 Notice of Allowance dated Aug. 14, 2017, issued in connection with U.S. Appl. No. 15/098,867, filed Apr. 14, 2016, 10 pages.  
 Notice of Allowance dated Jun. 14, 2017, issued in connection with U.S. Appl. No. 15/282,554, filed Sep. 30, 2016, 11 pages.  
 Notice of Allowance dated Aug. 16, 2017, issued in connection with U.S. Appl. No. 15/098,892, filed Apr. 14, 2016, 9 pages.  
 Notice of Allowance dated Aug. 17, 2017, issued in connection with U.S. Appl. No. 15/131,244, filed Apr. 18, 2016, 9 pages.  
 Notice of Allowance dated Aug. 22, 2017, issued in connection with U.S. Appl. No. 15/273,679, filed Sep. 22, 2016, 5 pages.  
 Non-Final Office Action dated Sep. 6, 2017, issued in connection with U.S. Appl. No. 15/131,254, filed Apr. 18, 2016, 13 pages.  
 Non-Final Office Action dated Sep. 14, 2017, issued in connection with U.S. Appl. No. 15/178,180, filed Jun. 9, 2016, 16 pages.  
 European Patent Office, European Extended Search Report dated Oct. 30, 2017, issued in connection with EP Application No. 17174435.2, 11 pages.  
 Final Office Action dated Oct. 6, 2017, issued in connection with U.S. Appl. No. 15/098,760, filed Apr. 14, 2016, 25 pages.  
 Final Office Action dated Apr. 13, 2018, issued in connection with U.S. Appl. No. 15/131,254, filed Apr. 18, 2016, 18 pages.  
 Final Office Action dated Apr. 13, 2018, issued in connection with U.S. Appl. No. 15/438,744, filed Feb. 21, 2017, 20 pages.  
 Fiorenza Arisio et al. "Deliverable 1.1 User Study, analysis of requirements and definition of the application task," May 31, 2012, [http://dirha.fbk.eu/sites/dirha.fbk.eu/files/docs/DIRHA\\_D1.1.](http://dirha.fbk.eu/sites/dirha.fbk.eu/files/docs/DIRHA_D1.1.), 31 pages.  
 Freiburger, Karl, "Development and Evaluation of Source Localization Algorithms for Coincident Microphone Arrays," Diploma Thesis, Apr. 1, 2010, 106 pages.  
 International Searching Authority, International Search Report and Written Opinion dated Nov. 22, 2017, issued in connection with International Application No. PCT/US2017/054063, filed on Sep. 28, 2017, 11 pages.  
 International Searching Authority, International Search Report and Written Opinion dated Oct. 23, 2017, issued in connection with International Application No. PCT/US2017/042170, filed on Jul. 14, 2017, 15 pages.  
 International Searching Authority, International Search Report and Written Opinion dated Oct. 24, 2017, issued in connection with International Application No. PCT/US2017/042227, filed on Jul. 14, 2017, 16 pages.  
 Morales-Cordovilla et al. "Room Localization for Distant Speech Recognition," Proceedings of Interspeech 2014, Sep. 14, 2014, 4 pages.  
 Non-Final Office Action dated Nov. 2, 2017, issued in connection with U.S. Appl. No. 15/584,782, filed May 2, 2017, 11 pages.  
 Non-Final Office Action dated Nov. 3, 2017, issued in connection with U.S. Appl. No. 15/438,741, filed Feb. 21, 2017, 11 pages.

(56)

**References Cited**

## OTHER PUBLICATIONS

Non-Final Office Action dated Jan. 10, 2018, issued in connection with U.S. Appl. No. 15/229,868, filed Aug. 5, 2016, 13 pages.  
Non-Final Office Action dated Jan. 10, 2018, issued in connection with U.S. Appl. No. 15/438,725, filed Feb. 21, 2017, 15 pages.  
Non-Final Office Action dated Mar. 16, 2018, issued in connection with U.S. Appl. No. 15/681,937, filed Aug. 21, 2017, 5 pages.  
Non-Final Office Action dated Apr. 18, 2018, issued in connection with U.S. Appl. No. 15/811,468, filed Nov. 13, 2017, 14 pages.  
Non-Final Office Action dated Feb. 20, 2018, issued in connection with U.S. Appl. No. 15/211,748, filed Jul. 15, 2016, 31 pages.  
Non-Final Office Action dated May 22, 2018, issued in connection with U.S. Appl. No. 15/946,599, filed Apr. 5, 2018, 19 pages.  
Non-Final Office Action dated Oct. 26, 2017, issued in connection with U.S. Appl. No. 15/438,744, filed Feb. 21, 2017, 12 pages.  
Non-Final Office Action dated Feb. 6, 2018, issued in connection with U.S. Appl. No. 15/211,689, filed Jul. 15, 2016, 32 pages.  
Non-Final Office Action dated Feb. 6, 2018, issued in connection with U.S. Appl. No. 15/237,133, filed Aug. 15, 2016, 6 pages.  
Non-Final Office Action dated Apr. 9, 2018, issued in connection with U.S. Appl. No. 15/804,776, filed Nov. 6, 2017, 18 pages.  
Non-Final Office Action dated May 9, 2018, issued in connection with U.S. Appl. No. 15/818,051, filed Nov. 20, 2017, 22 pages.  
Notice of Allowance dated Dec. 4, 2017, issued in connection with U.S. Appl. No. 15/277,810, filed Sep. 27, 2016, 5 pages.

Notice of Allowance dated Apr. 11, 2018, issued in connection with U.S. Appl. No. 15/719,454, filed Sep. 28, 2017, 15 pages.  
Notice of Allowance dated Dec. 13, 2017, issued in connection with U.S. Appl. No. 15/784,952, filed Oct. 16, 2017, 9 pages.  
Notice of Allowance dated Dec. 15, 2017, issued in connection with U.S. Appl. No. 15/223,218, filed Jul. 29, 2016, 7 pages.  
Notice of Allowance dated Mar. 20, 2018, issued in connection with U.S. Appl. No. 15/784,952, filed Oct. 16, 2017, 7 pages.  
Notice of Allowance dated Jan. 22, 2018, issued in connection with U.S. Appl. No. 15/178,180, filed Jun. 9, 2016, 9 pages.  
Notice of Allowance dated Dec. 29, 2017, issued in connection with U.S. Appl. No. 15/131,776, filed Apr. 18, 2016, 13 pages.  
Notice of Allowance dated Mar. 9, 2018, issued in connection with U.S. Appl. No. 15/584,782, filed May 2, 2017, 8 pages.  
Tsiami et al. "Experiments in acoustic source localization using sparse arrays in adverse indoors environments", 2014 22nd European Signal Processing Conference, Sep. 1, 2014, 5 pages.  
Vacher et al. "Recognition of voice commands by multisource ASR and noise cancellation in a smart home environment" Signal Processing Conference 2012 Proceedings of the 20th European, IEEE, Aug. 27, 2012, 5 pages.  
Xiao et al. "A Learning-Based Approach to Direction of Arrival Estimation in Noisy and Reverberant Environments," 2015 IEEE International Conference on Acoustics, Speech and Signal Processing, Apr. 19, 2015, 5 pages.

\* cited by examiner

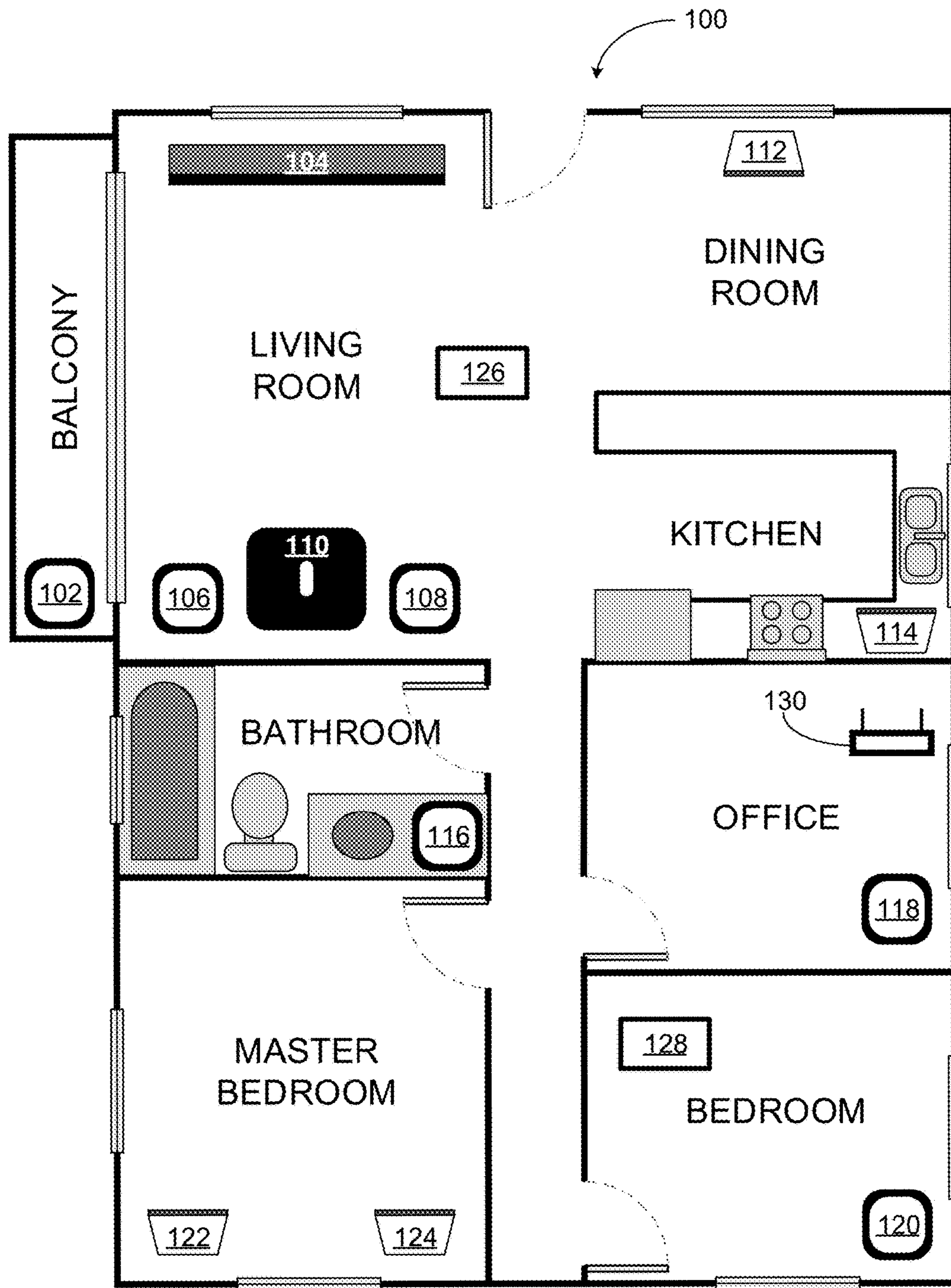


FIGURE 1

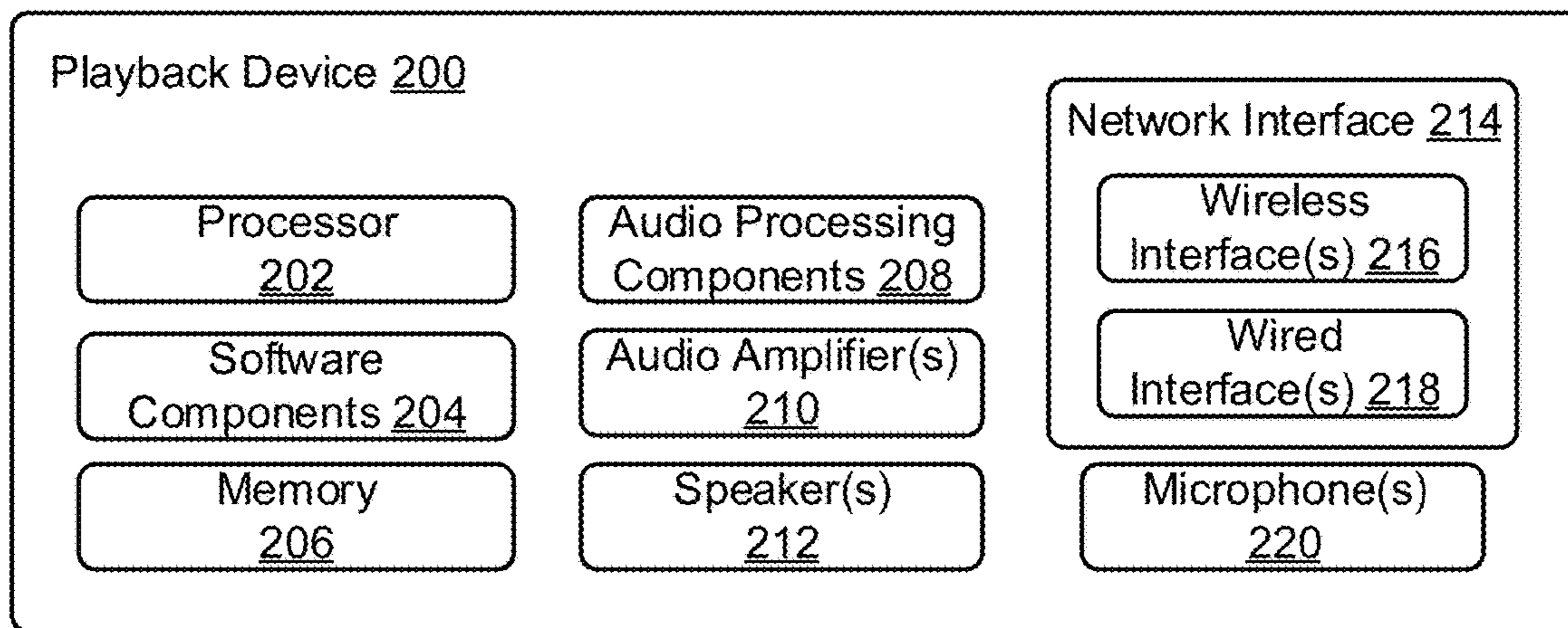


FIGURE 2

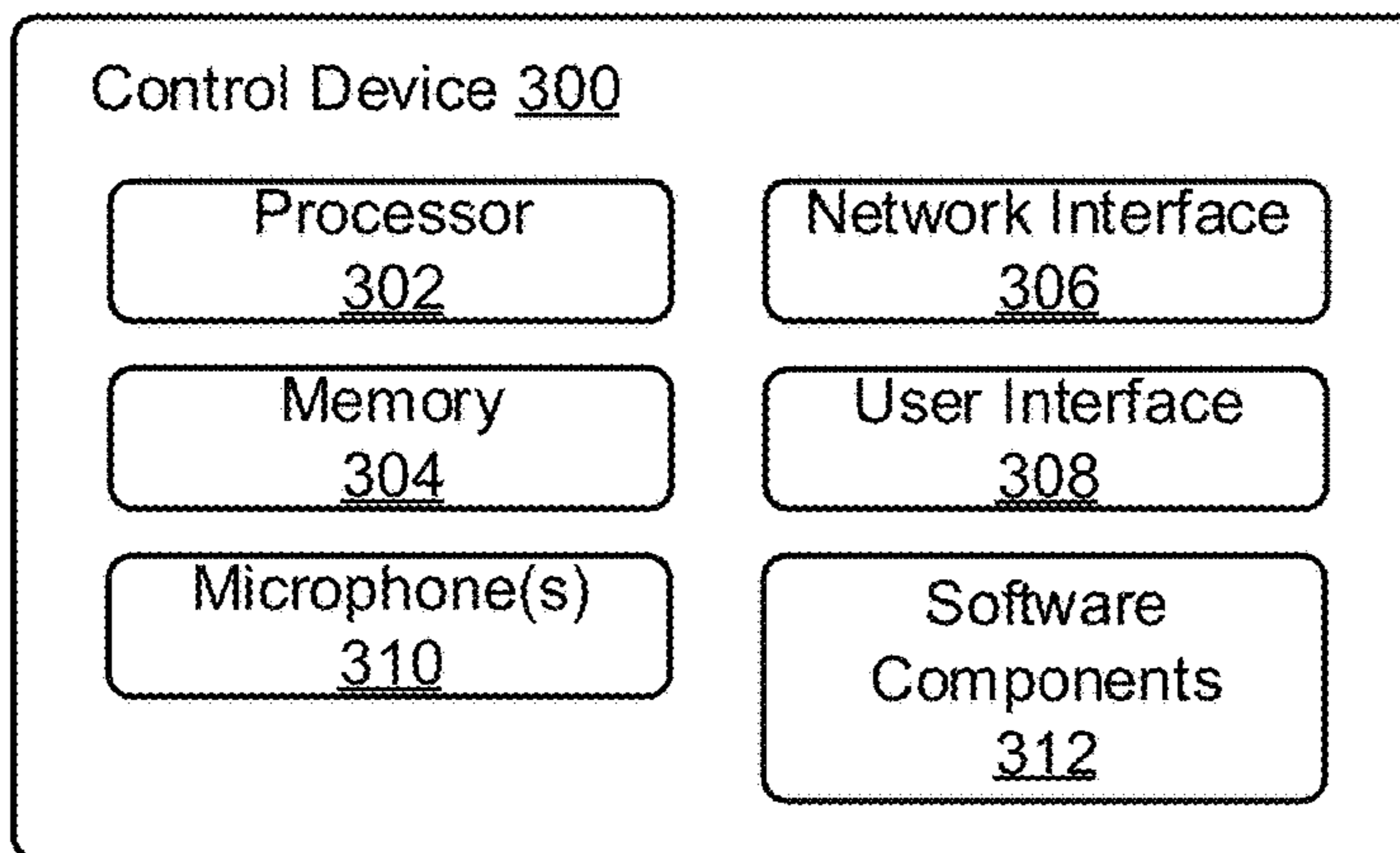


FIGURE 3

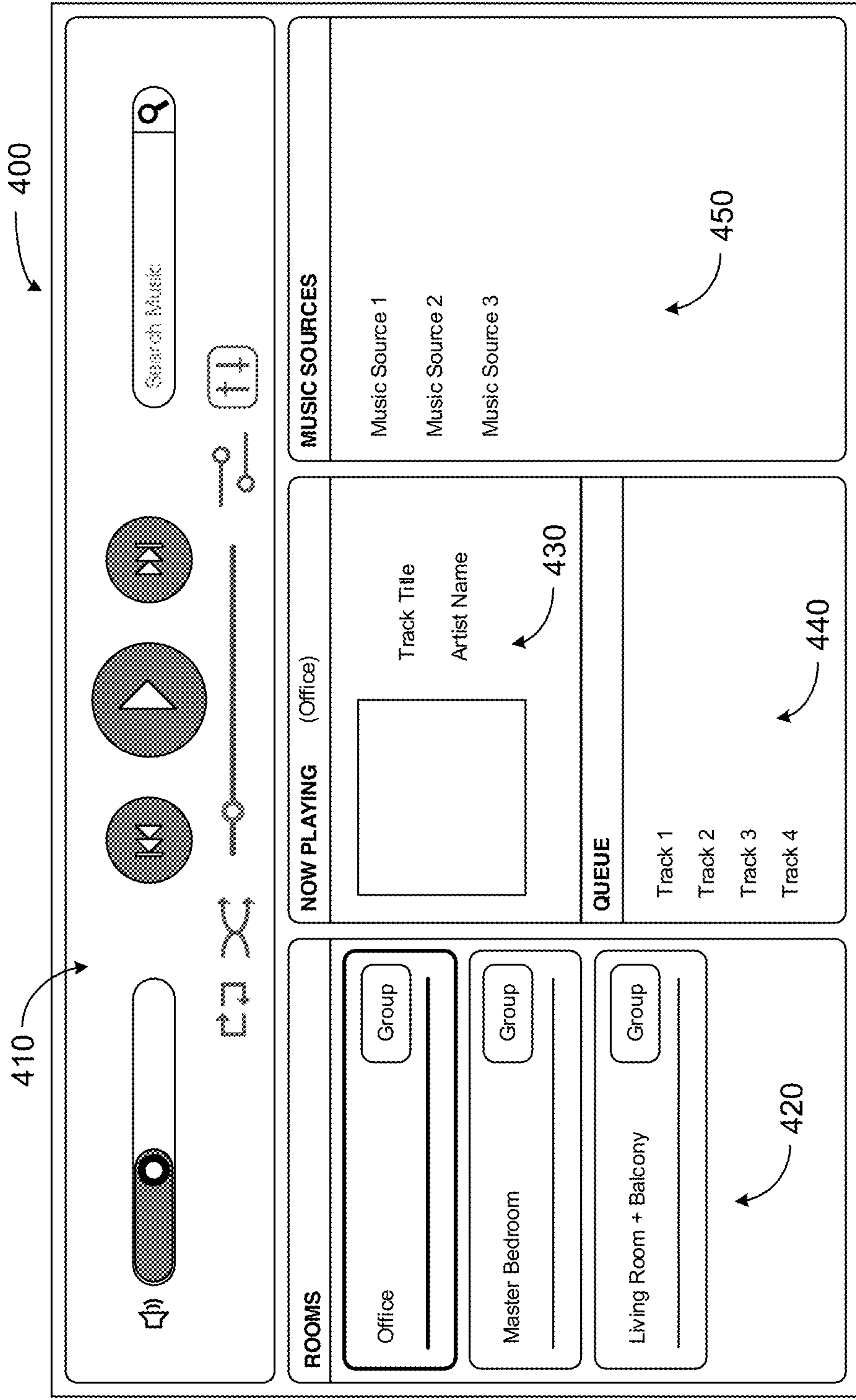


FIGURE 4



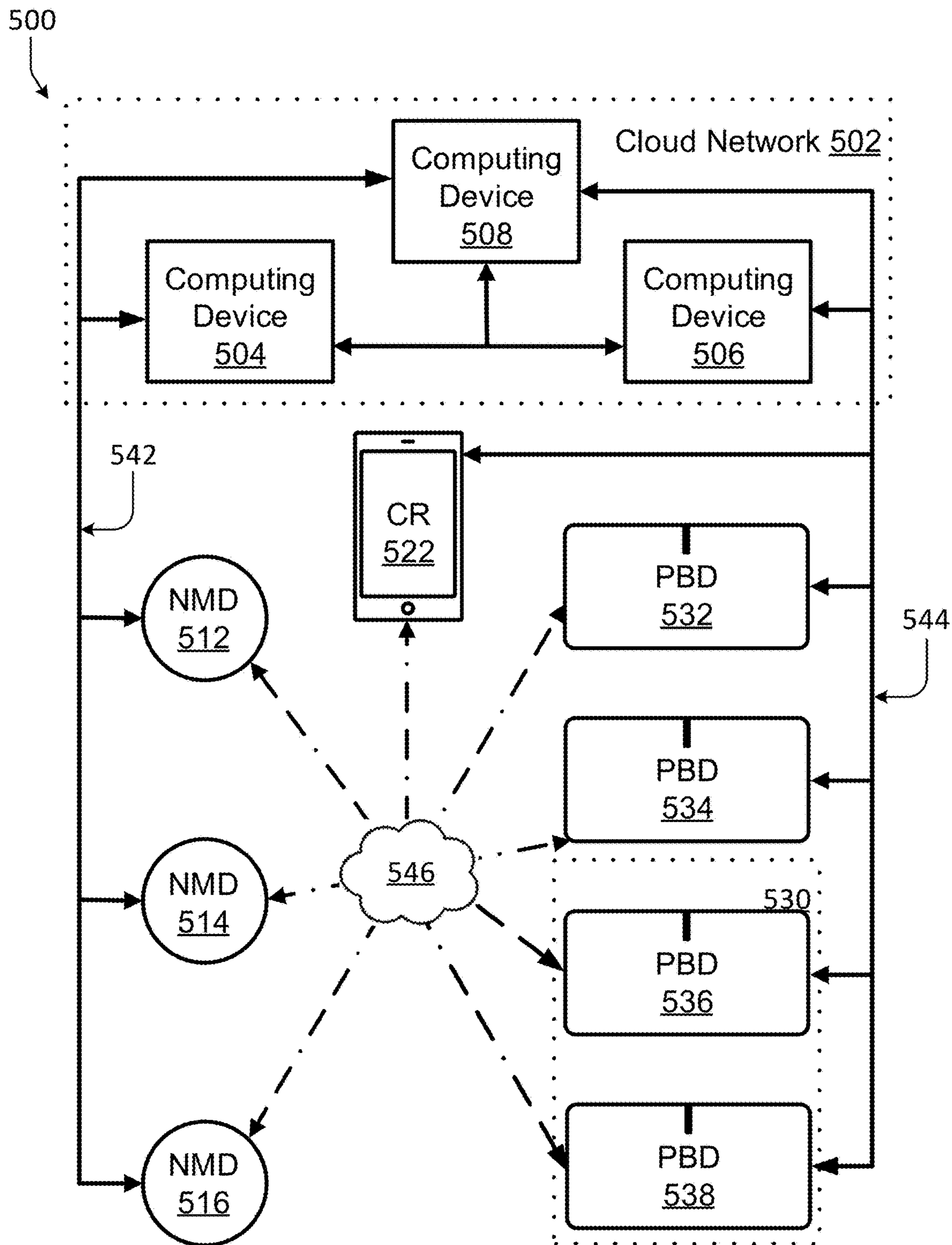


FIGURE 5

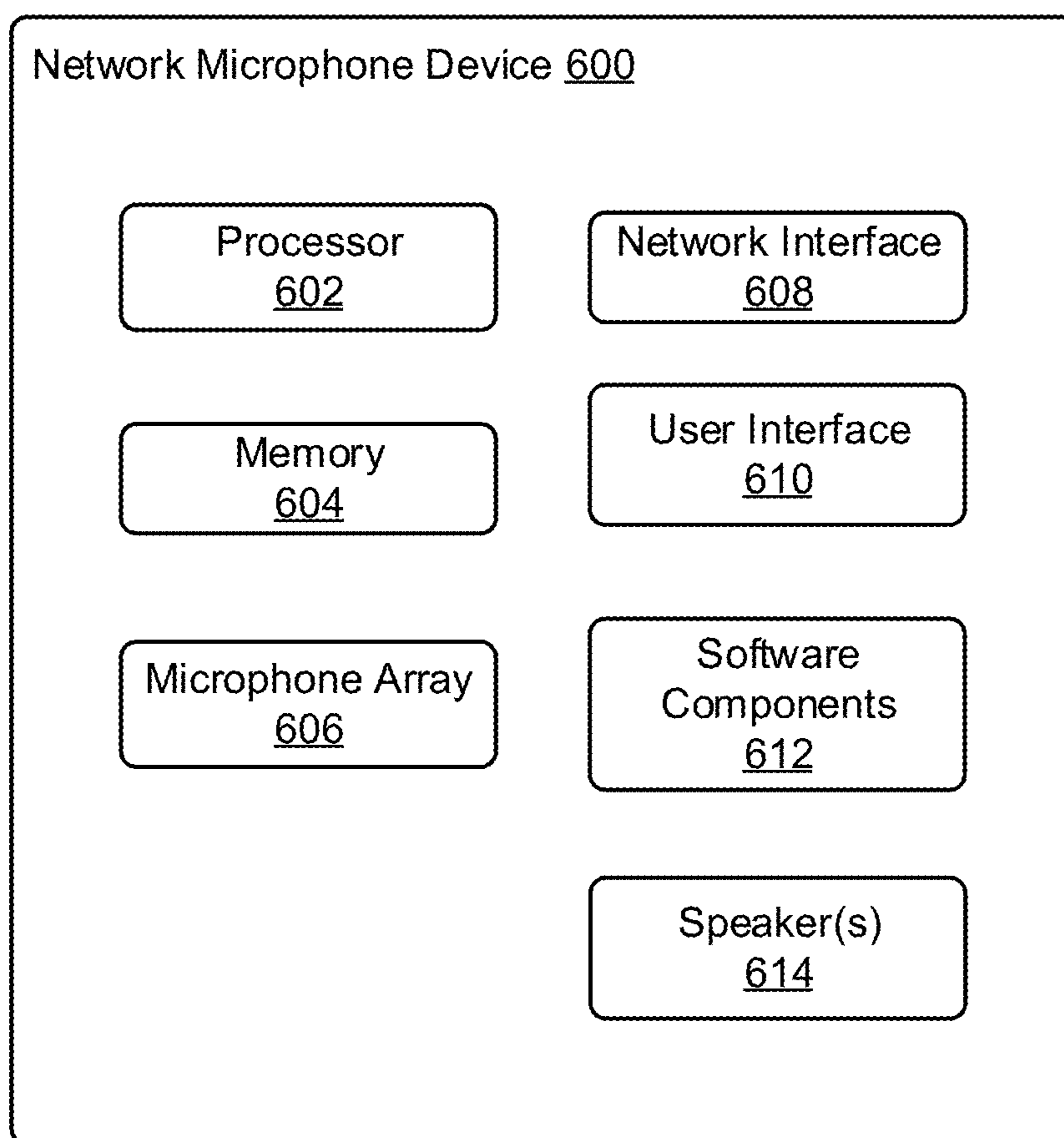


FIGURE 6

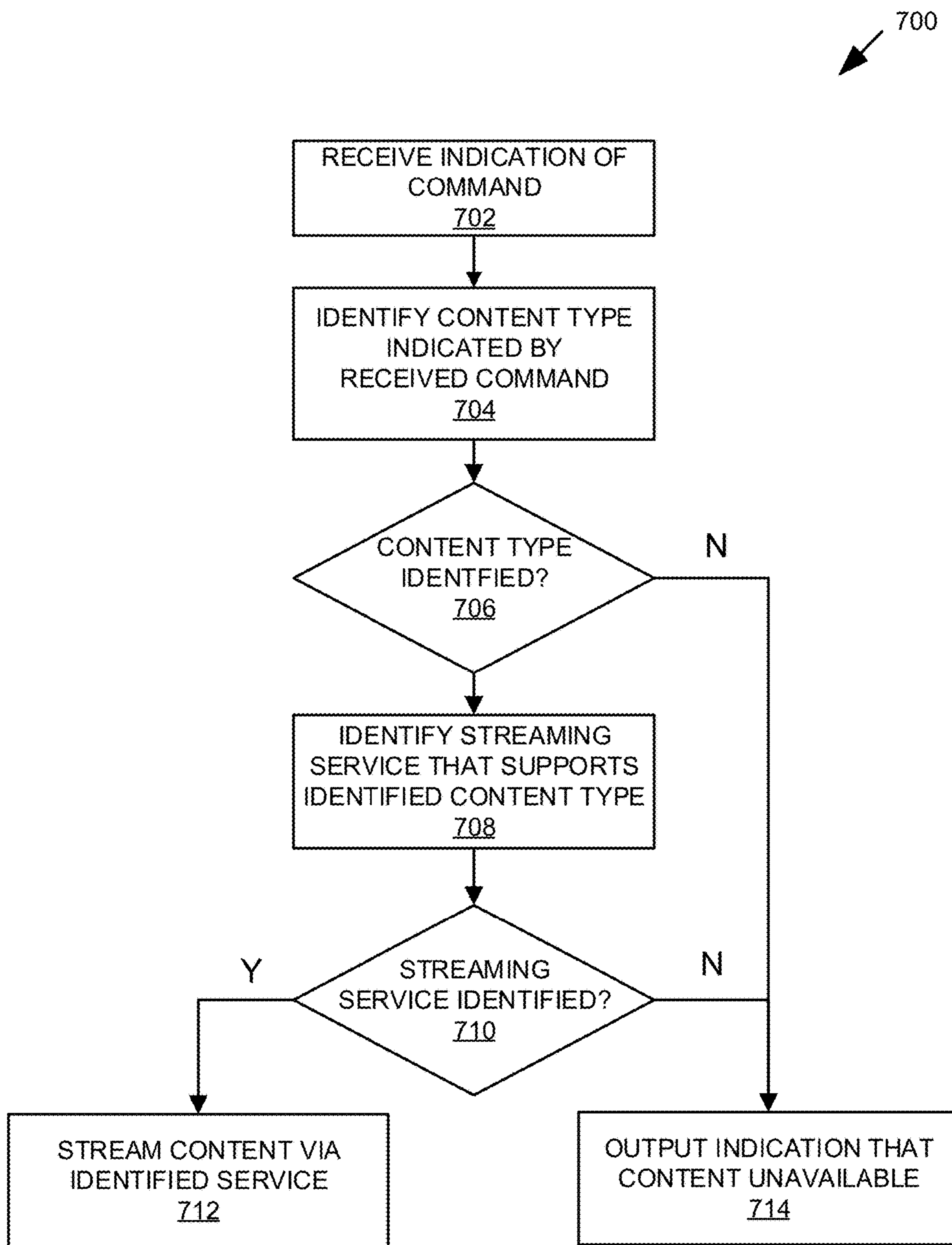


FIGURE 7

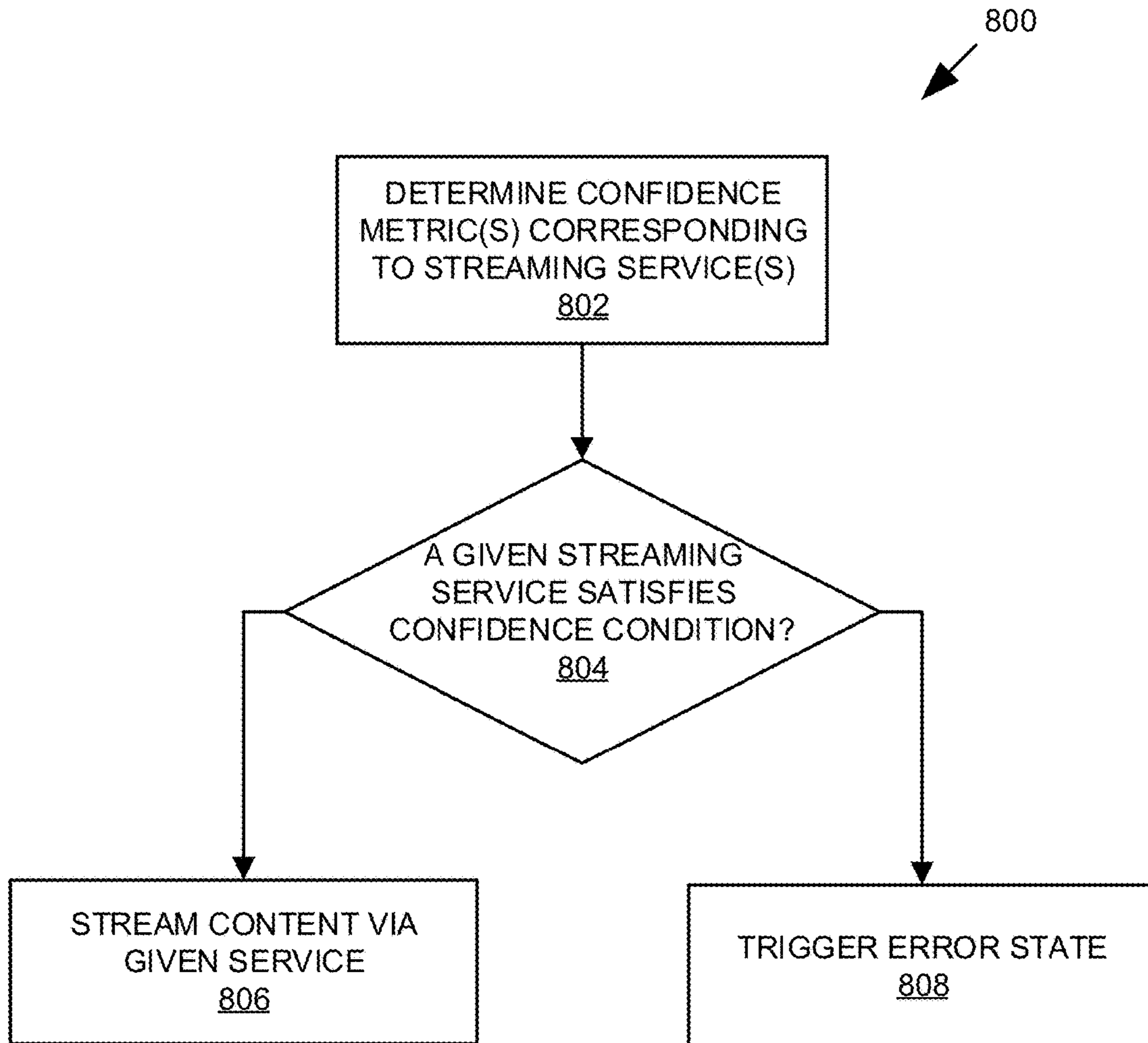


FIGURE 8

**1****MUSIC SERVICE SELECTION****CROSS-REFERENCE TO RELATED APPLICATIONS**

This application claims the benefit of priority under 35 USC § 119(e) to U.S. Provisional Application Ser. No. 62/298,410 filed on Feb. 22, 2016 and entitled "DEFAULT PLAYBACK DEVICE(S)", U.S. Provisional Application Ser. No. 62/298,418 filed on Feb. 22, 2016 and entitled "AUDIO RESPONSE PLAYBACK", U.S. Provisional Application Ser. No. 62/298,433 filed on Feb. 22, 2016 and entitled "ROOM-CORRECTED VOICE DETECTION", U.S. Provisional Application Ser. No. 62/298,439 filed on Feb. 22, 2016 and entitled "CONTENT MIXING", U.S. Provisional Application Ser. No. 62/298,425 filed on Feb. 22, 2016 and entitled "MUSIC SERVICE SELECTION", U.S. Provisional Application Ser. No. 62/298,350 filed on Feb. 22, 2016 and entitled "METADATA EXCHANGE INVOLVING A NETWORKED PLAYBACK SYSTEM AND A NETWORKED MICROPHONE SYSTEM", U.S. Provisional Application Ser. No. 62/298,388 filed on Feb. 22, 2016 and entitled "HANDLING OF LOSS OF PAIRING BETWEEN NETWORKED DEVICES", U.S. Provisional Application Ser. No. 62/298,393 filed on Feb. 22, 2016 and entitled "ACTION BASED ON USER ID", U.S. Provisional Application Ser. No. 62/312,350 filed on Mar. 23, 2016 and entitled "VOICE CONTROL OF A MEDIA PLAYBACK SYSTEM", the contents each of which are herein incorporated by reference in their entireties.

**FIELD OF THE DISCLOSURE**

The disclosure is related to consumer goods and, more particularly, to methods, systems, products, features, services, and other elements directed to media playback or some aspect thereof.

**BACKGROUND**

Options for accessing and listening to digital audio in an out-loud setting were limited until in 2003, when SONOS, Inc. filed for one of its first patent applications, entitled "Method for Synchronizing Audio Playback between Multiple Networked Devices," and began offering a media playback system for sale in 2005. The Sonos Wireless HiFi System enables people to experience music from many sources via one or more networked playback devices. Through a software control application installed on a smartphone, tablet, or computer, one can play what he or she wants in any room that has a networked playback device. Additionally, using the controller, for example, different songs can be streamed to each room with a playback device, rooms can be grouped together for synchronous playback, or the same song can be heard in all rooms synchronously.

Given the ever growing interest in digital media, there continues to be a need to develop consumer-accessible technologies to further enhance the listening experience.

**BRIEF DESCRIPTION OF THE DRAWINGS**

Features, aspects, and advantages of the presently disclosed technology may be better understood with regard to the following description, appended claims, and accompanying drawings where:

FIG. 1 shows an example media playback system configuration in which certain embodiments may be practiced;

**2**

FIG. 2 shows a functional block diagram of an example playback device;

FIG. 3 shows a functional block diagram of an example control device;

5 FIG. 4 shows an example controller interface;

FIG. 5 shows an example plurality of network devices;

FIG. 6 shows a function block diagram of an example network microphone device;

10 FIG. 7 is an example flow diagram related to identifying a streaming music service via a network microphone device;

FIG. 8 is another example flow diagram related to identifying a streaming music service via a network microphone device;

15 The drawings are for the purpose of illustrating example embodiments, but it is understood that the inventions are not limited to the arrangements and instrumentality shown in the drawings.

**DETAILED DESCRIPTION****I. Overview**

Embodiments described herein relate to identifying and accessing suitable streaming services (e.g. streaming audio tracks) based on commands.

25 Streaming services may be identified and accessed based upon voice commands provided by a network microphone device. The network microphone device may be a device which receives speech via a microphone and performs a function based upon the the speech. As an example, the microphone network may receive a voice command "play Pandora® 70's rock radio," determine that the speech is a command play a specific station from a specific streaming service, and then facilitate the playback of the station from that service. In other implementations, the selection of streaming service may be based upon commands input (e.g. textual input) via a user interface of a controller device. Other types of commands are also possible.

30 In another example, the commands may not explicitly specify which streaming service a user desires to be accessed. For example, a user may issue more generic commands such as "play Queen" or "play 70's rock." In such an instance, existing systems are limited in their ability to distinguish what content type a user desires to be played (i.e. the artist or the album Queen, a 70's rock radio station or a "best of" album.) Additionally, such systems are unable to match a user's intended content type to the content type capabilities of the various streaming services (i.e. which services are capable of playing radio station, artists, albums.)

35 Given the ever increasing amount of content and number of streaming services available there is a need for a system that is capable of intelligently selecting the desired streaming service that matches the content type to provide to the user based upon the user command.

40 In embodiments, selection of the streaming services may be accomplished by a network configuration including a plurality of network capable devices. The network configuration may include network microphone devices, playback devices, computing devices and/or controller devices (e.g. tablet, smartphone) receiving, processing, and analyzing commands. The configuration may further involve retrieving and/or requesting audio content from one or more music content servers based on the processing and analysis of the received commands. The audio content then may then be obtained by a network microphone device, controller, and/or any number of playback devices to provide an audio playback experience based on the commands. The network configuration may take other forms as well.

In example implementations, selection of a streaming service based on a command may be determined by a number of criteria, individually or in combination. In one example, the selection of a streaming service may be dependent on the content type (e.g. song, genre, radio station) indicated by the command and whether a particular streaming service supports the indicated content type. In such a case, content type logic may be utilized to correlate the command or portions of the command to content types. The content type(s) identified via the logic may then be mapped to a streaming services having the available content types. Analysis of content types indicated via a command may be performed in various other ways.

In another instance, the selection of a streaming service may be based in part upon a user history which may take into account a user's streaming service preferences. The user preferences may be on a per zone basis, content type most played, among various other. Additionally, various forms of "external" data may be incorporated, including but not limited to, geographic, demographic, and weather type data. Other types of selection influencing criteria may exist.

In addition to selecting a music service to provide audio, the processing of the user command may cause alternate indications to be output. In such examples, the system may output a suggestion of a streaming service capable of playing a content type indicated by a command. In another example, the system may output an indication that "the content is unavailable." Such indications may be output at a network microphone device, controller, or at one or more playback devices via the network microphone device or controller.

In another aspect, the various selection criteria may serve as inputs of an algorithm to determine confidence metrics for various streaming services. A confidence metric may be an indication whether the particular music service is what the user may desire to listen to. For example, the confidence level may be a probability value or percentage (e.g. 1-100) assigned to streaming services. In one example, the streaming service with the highest confidence metric may be provided for streaming. In another instance, an error state may be triggered if the highest calculated confidence metric does not exceed a threshold confidence value or if the top N confidence levels are within a specified range of one another. In such a case, an error state may cause the network microphone device to (1) output an indication that "the content is unavailable," (2) ask the user to repeat/further specify the command, (3) cause audio to be played by a preferred partner or default service, among other possibilities. The confidence metric may be used in a variety of many other manners.

While some examples described herein may refer to functions performed by given actors such as "users" and/or other entities, it should be understood that this is for purposes of explanation only. The claims should not be interpreted to require action by any such example actor unless explicitly required by the language of the claims themselves. It will be understood by one of ordinary skill in the art that this disclosure includes numerous other embodiments.

## II. Example Operating Environment

FIG. 1 shows an example configuration of a media playback system 100 in which one or more embodiments disclosed herein may be practiced or implemented. The media playback system 100 as shown is associated with an example home environment having several rooms and spaces, such as for example, a master bedroom, an office, a dining room, and a living room. As shown in the example of FIG. 1, the media playback system 100 includes playback

devices 102-124, control devices 126 and 128, and a wired or wireless network router 130.

Further discussions relating to the different components of the example media playback system 100 and how the different components may interact to provide a user with a media experience may be found in the following sections. While discussions herein may generally refer to the example media playback system 100, technologies described herein are not limited to applications within, among other things, the home environment as shown in FIG. 1. For instance, the technologies described herein may be useful in environments where multi-zone audio may be desired, such as, for example, a commercial setting like a restaurant, mall or airport, a vehicle like a sports utility vehicle (SUV), bus or car, a ship or boat, an airplane, and so on.

### a. Example Playback Devices

FIG. 2 shows a functional block diagram of an example playback device 200 that may be configured to be one or more of the playback devices 102-124 of the media playback system 100 of FIG. 1. The playback device 200 may include a processor 202, software components 204, memory 206, audio processing components 208, audio amplifier(s) 210, speaker(s) 212, a network interface 214 including wireless interface(s) 216 and wired interface(s) 218, and microphone(s) 220. In one case, the playback device 200 may not include the speaker(s) 212, but rather a speaker interface for connecting the playback device 200 to external speakers. In another case, the playback device 200 may include neither the speaker(s) 212 nor the audio amplifier(s) 210, but rather an audio interface for connecting the playback device 200 to an external audio amplifier or audio-visual receiver.

In one example, the processor 202 may be a clock-driven computing component configured to process input data according to instructions stored in the memory 206. The memory 206 may be a tangible computer-readable medium configured to store instructions executable by the processor 202. For instance, the memory 206 may be data storage that can be loaded with one or more of the software components 204 executable by the processor 202 to achieve certain functions. In one example, the functions may involve the playback device 200 retrieving audio data from an audio source or another playback device. In another example, the functions may involve the playback device 200 sending audio data to another device or playback device on a network. In yet another example, the functions may involve pairing of the playback device 200 with one or more playback devices to create a multi-channel audio environment.

Certain functions may involve the playback device 200 synchronizing playback of audio content with one or more other playback devices. During synchronous playback, a listener will preferably not be able to perceive time-delay differences between playback of the audio content by the playback device 200 and the one or more other playback devices. U.S. Pat. No. 8,234,395 entitled, "System and method for synchronizing operations among a plurality of independently clocked digital data processing devices," which is hereby incorporated by reference, provides in more detail some examples for audio playback synchronization among playback devices.

The memory 206 may further be configured to store data associated with the playback device 200, such as one or more zones and/or zone groups the playback device 200 is a part of, audio sources accessible by the playback device 200, or a playback queue that the playback device 200 (or some other playback device) may be associated with. The data may be stored as one or more state variables that are

periodically updated and used to describe the state of the playback device **200**. The memory **206** may also include the data associated with the state of the other devices of the media system, and shared from time to time among the devices so that one or more of the devices have the most recent data associated with the system. Other embodiments are also possible.

The audio processing components **208** may include one or more digital-to-analog converters (DAC), an audio pre-processing component, an audio enhancement component or a digital signal processor (DSP), and so on. In one embodiment, one or more of the audio processing components **208** may be a subcomponent of the processor **202**. In one example, audio content may be processed and/or intentionally altered by the audio processing components **208** to produce audio signals. The produced audio signals may then be provided to the audio amplifier(s) **210** for amplification and playback through speaker(s) **212**. Particularly, the audio amplifier(s) **210** may include devices configured to amplify audio signals to a level for driving one or more of the speakers **212**. The speaker(s) **212** may include an individual transducer (e.g., a “driver”) or a complete speaker system involving an enclosure with one or more drivers. A particular driver of the speaker(s) **212** may include, for example, a subwoofer (e.g., for low frequencies), a mid-range driver (e.g., for middle frequencies), and/or a tweeter (e.g., for high frequencies). In some cases, each transducer in the one or more speakers **212** may be driven by an individual corresponding audio amplifier of the audio amplifier(s) **210**. In addition to producing analog signals for playback by the playback device **200**, the audio processing components **208** may be configured to process audio content to be sent to one or more other playback devices for playback.

Audio content to be processed and/or played back by the playback device **200** may be received from an external source, such as via an audio line-in input connection (e.g., an auto-detecting 3.5 mm audio line-in connection) or the network interface **214**.

The network interface **214** may be configured to facilitate a data flow between the playback device **200** and one or more other devices on a data network. As such, the playback device **200** may be configured to receive audio content over the data network from one or more other playback devices in communication with the playback device **200**, network devices within a local area network, or audio content sources over a wide area network such as the Internet. In one example, the audio content and other signals transmitted and received by the playback device **200** may be transmitted in the form of digital packet data containing an Internet Protocol (IP)-based source address and IP-based destination addresses. In such a case, the network interface **214** may be configured to parse the digital packet data such that the data destined for the playback device **200** is properly received and processed by the playback device **200**.

As shown, the network interface **214** may include wireless interface(s) **216** and wired interface(s) **218**. The wireless interface(s) **216** may provide network interface functions for the playback device **200** to wirelessly communicate with other devices (e.g., other playback device(s), speaker(s), receiver(s), network device(s), control device(s) within a data network the playback device **200** is associated with) in accordance with a communication protocol (e.g., any wireless standard including IEEE 802.11a, 802.11b, 802.11g, 802.11n, 802.11ac, 802.15, 4G mobile communication standard, and so on). The wired interface(s) **218** may provide network interface functions for the playback device **200** to communicate over a wired connection with other devices in

accordance with a communication protocol (e.g., IEEE 802.3). While the network interface **214** shown in FIG. 2 includes both wireless interface(s) **216** and wired interface(s) **218**, the network interface **214** may in some embodiments include only wireless interface(s) or only wired interface(s).

The microphone(s) **220** may be arranged to detect sound in the environment of the playback device **200**. For instance, the microphone(s) may be mounted on an exterior wall of a housing of the playback device. The microphone(s) may be any type of microphone now known or later developed such as a condenser microphone, electret condenser microphone, or a dynamic microphone. The microphone(s) may be sensitive to a portion of the frequency range of the speaker(s) **220**. One or more of the speaker(s) **220** may operate in reverse as the microphone(s) **220**. In some aspects, the playback device **200** might not have microphone(s) **220**.

In one example, the playback device **200** and one other playback device may be paired to play two separate audio components of audio content. For instance, playback device **200** may be configured to play a left channel audio component, while the other playback device may be configured to play a right channel audio component, thereby producing or enhancing a stereo effect of the audio content. The paired playback devices (also referred to as “bonded playback devices”) may further play audio content in synchrony with other playback devices.

In another example, the playback device **200** may be sonically consolidated with one or more other playback devices to form a single, consolidated playback device. A consolidated playback device may be configured to process and reproduce sound differently than an unconsolidated playback device or playback devices that are paired, because a consolidated playback device may have additional speaker drivers through which audio content may be rendered. For instance, if the playback device **200** is a playback device designed to render low frequency range audio content (i.e. a subwoofer), the playback device **200** may be consolidated with a playback device designed to render full frequency range audio content. In such a case, the full frequency range playback device, when consolidated with the low frequency playback device **200**, may be configured to render only the mid and high frequency components of audio content, while the low frequency range playback device **200** renders the low frequency component of the audio content. The consolidated playback device may further be paired with a single playback device or yet another consolidated playback device.

By way of illustration, SONOS, Inc. presently offers (or has offered) for sale certain playback devices including a “PLAY:1,” “PLAY:3,” “PLAY:5,” “PLAYBAR,” “CONNECT:AMP,” “CONNECT,” and “SUB.” Any other past, present, and/or future playback devices may additionally or alternatively be used to implement the playback devices of example embodiments disclosed herein. Additionally, it is understood that a playback device is not limited to the example illustrated in FIG. 2 or to the SONOS product offerings. For example, a playback device may include a wired or wireless headphone. In another example, a playback device may include or interact with a docking station for personal mobile media playback devices. In yet another example, a playback device may be integral to another device or component such as a television, a lighting fixture, or some other device for indoor or outdoor use.

#### b. Example Playback Zone Configurations

Referring back to the media playback system **100** of FIG. 1, the environment may have one or more playback zones,

each with one or more playback devices. The media playback system 100 may be established with one or more playback zones, after which one or more zones may be added, or removed to arrive at the example configuration shown in FIG. 1. Each zone may be given a name according

to a different room or space such as an office, bathroom, master bedroom, bedroom, kitchen, dining room, living room, and/or balcony. In one case, a single playback zone may include multiple rooms or spaces. In another case, a single room or space may include multiple playback zones. As shown in FIG. 1, the balcony, dining room, kitchen, bathroom, office, and bedroom zones each have one playback device, while the living room and master bedroom zones each have multiple playback devices. In the living room zone, playback devices 104, 106, 108, and 110 may be configured to play audio content in synchrony as individual playback devices, as one or more bonded playback devices, as one or more consolidated playback devices, or any combination thereof. Similarly, in the case of the master bedroom, playback devices 122 and 124 may be configured to play audio content in synchrony as individual playback devices, as a bonded playback device, or as a consolidated playback device.

In one example, one or more playback zones in the environment of FIG. 1 may each be playing different audio content. For instance, the user may be grilling in the balcony zone and listening to hip hop music being played by the playback device 102 while another user may be preparing food in the kitchen zone and listening to classical music being played by the playback device 114. In another example, a playback zone may play the same audio content in synchrony with another playback zone. For instance, the user may be in the office zone where the playback device 118 is playing the same rock music that is being played by playback device 102 in the balcony zone. In such a case, playback devices 102 and 118 may be playing the rock music in synchrony such that the user may seamlessly (or at least substantially seamlessly) enjoy the audio content that is being played out-loud while moving between different playback zones. Synchronization among playback zones may be achieved in a manner similar to that of synchronization among playback devices, as described in previously referenced U.S. Pat. No. 8,234,395.

As suggested above, the zone configurations of the media playback system 100 may be dynamically modified, and in some embodiments, the media playback system 100 supports numerous configurations. For instance, if a user physically moves one or more playback devices to or from a zone, the media playback system 100 may be reconfigured to accommodate the change(s). For instance, if the user physically moves the playback device 102 from the balcony zone to the office zone, the office zone may now include both the playback device 118 and the playback device 102. The playback device 102 may be paired or grouped with the office zone and/or renamed if so desired via a control device such as the control devices 126 and 128. On the other hand, if the one or more playback devices are moved to a particular area in the home environment that is not already a playback zone, a new playback zone may be created for the particular area.

Further, different playback zones of the media playback system 100 may be dynamically combined into zone groups or split up into individual playback zones. For instance, the dining room zone and the kitchen zone 114 may be combined into a zone group for a dinner party such that playback devices 112 and 114 may render audio content in synchrony. On the other hand, the living room zone may be split into a

television zone including playback device 104, and a listening zone including playback devices 106, 108, and 110, if the user wishes to listen to music in the living room space while another user wishes to watch television.

#### c. Example Control Devices

FIG. 3 shows a functional block diagram of an example control device 300 that may be configured to be one or both of the control devices 126 and 128 of the media playback system 100. As shown, the control device 300 may include a processor 302, memory 304, a network interface 306, a user interface 308, microphone(s) 310, and software components 312. In one example, the control device 300 may be a dedicated controller for the media playback system 100. In another example, the control device 300 may be a network device on which media playback system controller application software may be installed, such as for example, an iPhone™, iPad™ or any other smart phone, tablet or network device (e.g., a networked computer such as a PC or Mac™).

The processor 302 may be configured to perform functions relevant to facilitating user access, control, and configuration of the media playback system 100. The memory 304 may be data storage that can be loaded with one or more of the software components executable by the processor 302 to perform those functions. The memory 304 may also be configured to store the media playback system controller application software and other data associated with the media playback system 100 and the user.

In one example, the network interface 306 may be based on an industry standard (e.g., infrared, radio, wired standards including IEEE 802.3, wireless standards including IEEE 802.11a, 802.11b, 802.11g, 802.11n, 802.11ac, 802.15, 4G mobile communication standard, and so on). The network interface 306 may provide a means for the control device 300 to communicate with other devices in the media playback system 100. In one example, data and information (e.g., such as a state variable) may be communicated between control device 300 and other devices via the network interface 306. For instance, playback zone and zone group configurations in the media playback system 100 may be received by the control device 300 from a playback device or another network device, or transmitted by the control device 300 to another playback device or network device via the network interface 306. In some cases, the other network device may be another control device.

Playback device control commands such as volume control and audio playback control may also be communicated from the control device 300 to a playback device via the network interface 306. As suggested above, changes to configurations of the media playback system 100 may also be performed by a user using the control device 300. The configuration changes may include adding/removing one or more playback devices to/from a zone, adding/removing one or more zones to/from a zone group, forming a bonded or consolidated player, separating one or more playback devices from a bonded or consolidated player, among others. Accordingly, the control device 300 may sometimes be referred to as a controller, whether the control device 300 is a dedicated controller or a network device on which media playback system controller application software is installed.

Control device 300 may include microphone(s) 310. Microphone(s) 310 may be arranged to detect sound in the environment of the control device 300. Microphone(s) 310 may be any type of microphone now known or later developed such as a condenser microphone, electret condenser microphone, or a dynamic microphone. The microphone(s) may be sensitive to a portion of a frequency range. Two or



more microphones **310** may be arranged to capture location information of an audio source (e.g., voice, audible sound) and/or to assist in filtering background noise.

The user interface **308** of the control device **300** may be configured to facilitate user access and control of the media playback system **100**, by providing a controller interface such as the controller interface **400** shown in FIG. 4. The controller interface **400** includes a playback control region **410**, a playback zone region **420**, a playback status region **430**, a playback queue region **440**, and an audio content sources region **450**. The user interface **400** as shown is just one example of a user interface that may be provided on a network device such as the control device **300** of FIG. 3 (and/or the control devices **126** and **128** of FIG. 1) and accessed by users to control a media playback system such as the media playback system **100**. Other user interfaces of varying formats, styles, and interactive sequences may alternatively be implemented on one or more network devices to provide comparable control access to a media playback system.

The playback control region **410** may include selectable (e.g., by way of touch or by using a cursor) icons to cause playback devices in a selected playback zone or zone group to play or pause, fast forward, rewind, skip to next, skip to previous, enter/exit shuffle mode, enter/exit repeat mode, enter/exit cross fade mode. The playback control region **410** may also include selectable icons to modify equalization settings, and playback volume, among other possibilities.

The playback zone region **420** may include representations of playback zones within the media playback system **100**. In some embodiments, the graphical representations of playback zones may be selectable to bring up additional selectable icons to manage or configure the playback zones in the media playback system, such as a creation of bonded zones, creation of zone groups, separation of zone groups, and renaming of zone groups, among other possibilities.

For example, as shown, a “group” icon may be provided within each of the graphical representations of playback zones. The “group” icon provided within a graphical representation of a particular zone may be selectable to bring up options to select one or more other zones in the media playback system to be grouped with the particular zone. Once grouped, playback devices in the zones that have been grouped with the particular zone will be configured to play audio content in synchrony with the playback device(s) in the particular zone. Analogously, a “group” icon may be provided within a graphical representation of a zone group. In this case, the “group” icon may be selectable to bring up options to deselect one or more zones in the zone group to be removed from the zone group. Other interactions and implementations for grouping and ungrouping zones via a user interface such as the user interface **400** are also possible. The representations of playback zones in the playback zone region **420** may be dynamically updated as playback zone or zone group configurations are modified.

The playback status region **430** may include graphical representations of audio content that is presently being played, previously played, or scheduled to play next in the selected playback zone or zone group. The selected playback zone or zone group may be visually distinguished on the user interface, such as within the playback zone region **420** and/or the playback status region **430**. The graphical representations may include track title, artist name, album name, album year, track length, and other relevant information that may be useful for the user to know when controlling the media playback system via the user interface **400**.

The playback queue region **440** may include graphical representations of audio content in a playback queue associated with the selected playback zone or zone group. In some embodiments, each playback zone or zone group may be associated with a playback queue containing information corresponding to zero or more audio items for playback by the playback zone or zone group. For instance, each audio item in the playback queue may comprise a uniform resource identifier (URI), a uniform resource locator (URL) or some other identifier that may be used by a playback device in the playback zone or zone group to find and/or retrieve the audio item from a local audio content source or a networked audio content source, possibly for playback by the playback device.

In one example, a playlist may be added to a playback queue, in which case information corresponding to each audio item in the playlist may be added to the playback queue. In another example, audio items in a playback queue may be saved as a playlist. In a further example, a playback queue may be empty, or populated but “not in use” when the playback zone or zone group is playing continuously streaming audio content, such as Internet radio that may continue to play until otherwise stopped, rather than discrete audio items that have playback durations. In an alternative embodiment, a playback queue can include Internet radio and/or other streaming audio content items and be “in use” when the playback zone or zone group is playing those items. Other examples are also possible.

When playback zones or zone groups are “grouped” or “ungrouped,” playback queues associated with the affected playback zones or zone groups may be cleared or re-associated. For example, if a first playback zone including a first playback queue is grouped with a second playback zone including a second playback queue, the established zone group may have an associated playback queue that is initially empty, that contains audio items from the first playback queue (such as if the second playback zone was added to the first playback zone), that contains audio items from the second playback queue (such as if the first playback zone was added to the second playback zone), or a combination of audio items from both the first and second playback queues. Subsequently, if the established zone group is ungrouped, the resulting first playback zone may be re-associated with the previous first playback queue, or be associated with a new playback queue that is empty or contains audio items from the playback queue associated with the established zone group before the established zone group was ungrouped. Similarly, the resulting second playback zone may be re-associated with the previous second playback queue, or be associated with a new playback queue that is empty, or contains audio items from the playback queue associated with the established zone group before the established zone group was ungrouped. Other examples are also possible.

Referring back to the user interface **400** of FIG. 4, the graphical representations of audio content in the playback queue region **440** may include track titles, artist names, track lengths, and other relevant information associated with the audio content in the playback queue. In one example, graphical representations of audio content may be selectable to bring up additional selectable icons to manage and/or manipulate the playback queue and/or audio content represented in the playback queue. For instance, a represented audio content may be removed from the playback queue, moved to a different position within the playback queue, or selected to be played immediately, or after any currently playing audio content, among other possibilities. A playback

queue associated with a playback zone or zone group may be stored in a memory on one or more playback devices in the playback zone or zone group, on a playback device that is not in the playback zone or zone group, and/or some other designated device.

The audio content sources region **450** may include graphical representations of selectable audio content sources from which audio content may be retrieved and played by the selected playback zone or zone group. Discussions pertaining to audio content sources may be found in the following section.

#### d. Example Audio Content Sources

As indicated previously, one or more playback devices in a zone or zone group may be configured to retrieve for playback audio content (e.g. according to a corresponding URI or URL for the audio content) from a variety of available audio content sources. In one example, audio content may be retrieved by a playback device directly from a corresponding audio content source (e.g., a line-in connection). In another example, audio content may be provided to a playback device over a network via one or more other playback devices or network devices.

Example audio content sources may include a memory of one or more playback devices in a media playback system such as the media playback system **100** of FIG. **1**, local music libraries on one or more network devices (such as a control device, a network-enabled personal computer, or a networked-attached storage (NAS), for example), streaming audio services providing audio content via the Internet (e.g., the cloud), or audio sources connected to the media playback system via a line-in input connection on a playback device or network device, among other possibilities.

In some embodiments, audio content sources may be regularly added or removed from a media playback system such as the media playback system **100** of FIG. **1**. In one example, an indexing of audio items may be performed whenever one or more audio content sources are added, removed or updated. Indexing of audio items may involve scanning for identifiable audio items in all folders/directory shared over a network accessible by playback devices in the media playback system, and generating or updating an audio content database containing metadata (e.g., title, artist, album, track length, among others) and other associated information, such as a URI or URL for each identifiable audio item found. Other examples for managing and maintaining audio content sources may also be possible.

The above discussions relating to playback devices, controller devices, playback zone configurations, and media content sources provide only some examples of operating environments within which functions and methods described below may be implemented. Other operating environments and configurations of media playback systems, playback devices, and network devices not explicitly described herein may also be applicable and suitable for implementation of the functions and methods.

#### e. Example Plurality of Networked Devices

FIG. **5** shows an example plurality of devices **500** that may be configured to provide an audio playback experience based on voice control. One having ordinary skill in the art will appreciate that the devices shown in FIG. **5** are for illustrative purposes only, and variations including different and/or additional devices may be possible. As shown, the plurality of devices **500** includes computing devices **504**, **506**, and **508**; network microphone devices (NMDs) **512**, **514**, and **516**; playback devices (PBDs) **532**, **534**, **536**, and **538**; and a controller device (CR) **522**.

Each of the plurality of devices **500** may be network-capable devices that can establish communication with one or more other devices in the plurality of devices according to one or more network protocols, such as NFC, Bluetooth, Ethernet, and IEEE 802.11, among other examples, over one or more types of networks, such as wide area networks (WAN), local area networks (LAN), and personal area networks (PAN), among other possibilities.

As shown, the computing devices **504**, **506**, and **508** may be part of a cloud network **502**. The cloud network **502** may include additional computing devices. In one example, the computing devices **504**, **506**, and **508** may be different servers. In another example, two or more of the computing devices **504**, **506**, and **508** may be modules of a single server. Analogously, each of the computing device **504**, **506**, and **508** may include one or more modules or servers. For ease of illustration purposes herein, each of the computing devices **504**, **506**, and **508** may be configured to perform particular functions within the cloud network **502**. For instance, computing device **508** may be a source of audio content for a streaming music service.

As shown, the computing device **504** may be configured to interface with NMDs **512**, **514**, and **516** via communication path **542**. NMDs **512**, **514**, and **516** may be components of one or more "Smart Home" systems. In one case, NMDs **512**, **514**, and **516** may be physically distributed throughout a household, similar to the distribution of devices shown in FIG. **1**. In another case, two or more of the NMDs **512**, **514**, and **516** may be physically positioned within relative close proximity of one another. Communication path **542** may comprise one or more types of networks, such as a WAN including the Internet, LAN, and/or PAN, among other possibilities.

In one example, one or more of the NMDs **512**, **514**, and **516** may be devices configured primarily for audio detection. In another example, one or more of the NMDs **512**, **514**, and **516** may be components of devices having various primary utilities. For instance, as discussed above in connection to FIGS. **2** and **3**, one or more of NMDs **512**, **514**, and **516** may be the microphone(s) **220** of playback device **200** or the microphone(s) **310** of network device **300**. Further, in some cases, one or more of NMDs **512**, **514**, and **516** may be the playback device **200** or network device **300**. In an example, one or more of NMDs **512**, **514**, and/or **516** may include multiple microphones arranged in a microphone array.

As shown, the computing device **506** may be configured to interface with CR **522** and PBDs **532**, **534**, **536**, and **538** via communication path **544**. In one example, CR **522** may be a network device such as the network device **200** of FIG. **2**. Accordingly, CR **522** may be configured to provide the controller interface **400** of FIG. **4**. Similarly, PBDs **532**, **534**, **536**, and **538** may be playback devices such as the playback device **300** of FIG. **3**. As such, PBDs **532**, **534**, **536**, and **538** may be physically distributed throughout a household as shown in FIG. **1**. For illustration purposes, PBDs **536** and **538** may be part of a bonded zone **530**, while PBDs **532** and **534** may be part of their own respective zones. As described above, the PBDs **532**, **534**, **536**, and **538** may be dynamically bonded, grouped, unbonded, and ungrouped. Communication path **544** may comprise one or more types of networks, such as a WAN including the Internet, LAN, and/or PAN, among other possibilities.

In one example, as with NMDs **512**, **514**, and **516**, CR **522** and PBDs **532**, **534**, **536**, and **538** may also be components of one or more "Smart Home" systems. In one case, PBDs **532**, **534**, **536**, and **538** may be distributed throughout the

same household as the NMDs **512**, **514**, and **516**. Further, as suggested above, one or more of PBDs **532**, **534**, **536**, and **538** may be one or more of NMDs **512**, **514**, and **516**.

The NMDs **512**, **514**, and **516** may be part of a local area network, and the communication path **542** may include an access point that links the local area network of the NMDs **512**, **514**, and **516** to the computing device **504** over a WAN (communication path not shown). Likewise, each of the NMDs **512**, **514**, and **516** may communicate with each other via such an access point.

Similarly, CR **522** and PBDs **532**, **534**, **536**, and **538** may be part of a local area network and/or a local playback network as discussed in previous sections, and the communication path **544** may include an access point that links the local area network and/or local playback network of CR **522** and PBDs **532**, **534**, **536**, and **538** to the computing device **506** over a WAN. As such, each of the CR **522** and PBDs **532**, **534**, **536**, and **538** may also communicate with each other over such an access point.

In one example, communication paths **542** and **544** may comprise the same access point. In an example, each of the NMDs **512**, **514**, and **516**, CR **522**, and PBDs **532**, **534**, **536**, and **538** may access the cloud network **502** via the same access point for a household.

As shown in FIG. **5**, each of the NMDs **512**, **514**, and **516**, CR **522**, and PBDs **532**, **534**, **536**, and **538** may also directly communicate with one or more of the other devices via communication means **546**. Communication means **546** as described herein may involve one or more forms of communication between the devices, according to one or more network protocols, over one or more types of networks, and/or may involve communication via one or more other network devices. For instance, communication means **546** may include one or more of for example, Bluetooth™ (IEEE 802.15), NFC, Wireless direct, and/or Proprietary wireless, among other possibilities.

In one example, CR **522** may communicate with NMD **512** over Bluetooth™, and communicate with PBD **534** over another local area network. In another example, NMD **514** may communicate with CR **522** over another local area network, and communicate with PBD **536** over Bluetooth. In a further example, each of the PBDs **532**, **534**, **536**, and **538** may communicate with each other according to a spanning tree protocol over a local playback network, while each communicating with CR **522** over a local area network, different from the local playback network. Other examples are also possible.

In some cases, communication means between the NMDs **512**, **514**, and **516**, CR **522**, and PBDs **532**, **534**, **536**, and **538** may change depending on types of communication between the devices, network conditions, and/or latency demands. For instance, communication means **546** may be used when NMD **516** is first introduced to the household with the PBDs **532**, **534**, **536**, and **538**. In one case, the NMD **516** may transmit identification information corresponding to the NMD **516** to PBD **538** via NFC, and PBD **538** may in response, transmit local area network information to NMD **516** via NFC (or some other form of communication). However, once NMD **516** has been configured within the household, communication means between NMD **516** and PBD **538** may change. For instance, NMD **516** may subsequently communicate with PBD **538** via communication path **542**, the cloud network **502**, and communication path **544**. In another example, the NMDs and PBDs may never communicate via local communications means **546**. In a

further example, the NMDs and PBDs may communicate primarily via local communications means **546**. Other examples are also possible.

In an illustrative example, NMDs **512**, **514**, and **516** may be configured to receive voice inputs to control PBDs **532**, **534**, **536**, and **538**. The available control commands may include any media playback system controls previously discussed, such as playback volume control, playback transport controls, music source selection, and grouping, among other possibilities. In one instance, NMD **512** may receive a voice input to control one or more of the PBDs **532**, **534**, **536**, and **538**. In response to receiving the voice input, NMD **512** may transmit via communication path **542**, the voice input to computing device **504** for processing. In one example, the computing device **504** may convert the voice input to an equivalent text command, and parse the text command to identify a command. Computing device **504** may then subsequently transmit the text command to the computing device **506**. In another example, the computing device **504** may convert the voice input to an equivalent text command, and then subsequently transmit the text command to the computing device **506**. The computing device **506** may then parse the text command to identify one or more playback commands.

For instance, if the text command is “Play ‘Track 1’ by ‘Artist 1’ from ‘Streaming Service 1’ in ‘Zone 1,’” The computing device **506** may identify (i) a URL for “Track 1” by “Artist 1” available from “Streaming Service 1,” and (ii) at least one playback device in “Zone 1.” In this example, the URL for “Track 1” by “Artist 1” from “Streaming Service 1” may be a URL pointing to computing device **508**, and “Zone 1” may be the bonded zone **530**. As such, upon identifying the URL and one or both of PBDs **536** and **538**, the computing device **506** may transmit via communication path **544** to one or both of PBDs **536** and **538**, the identified URL for playback. One or both of PBDs **536** and **538** may responsively retrieve audio content from the computing device **508** according to the received URL, and begin playing “Track 1” by “Artist 1” from “Streaming Service 1.”

One having ordinary skill in the art will appreciate that the above is just one illustrative example, and that other implementations are also possible. In one case, operations performed by one or more of the plurality of devices **500**, as described above, may be performed by one or more other devices in the plurality of device **500**. For instance, the conversion from voice input to the text command may be alternatively, partially, or wholly performed by another device or devices, such as NMD **512**, computing device **506**, PBD **536**, and/or PBD **538**. Analogously, the identification of the URL may be alternatively, partially, or wholly performed by another device or devices, such as NMD **512**, computing device **504**, PBD **536**, and/or PBD **538**.

#### f. Example Network Microphone Device

FIG. **6** shows a function block diagram of an example network microphone device **600** that may be configured to be one or more of NMDs **512**, **514**, and **516** of FIG. **5**. As shown, the network microphone device **600** includes a processor **602**, memory **604**, a microphone array **606**, a network interface **608**, a user interface **610**, software components **612**, and speaker(s) **614**. One having ordinary skill in the art will appreciate that other network microphone device configurations and arrangements are also possible. For instance, network microphone devices may alternatively exclude the speaker(s) **614** or have a single microphone instead of microphone array **606**.

The processor **602** may include one or more processors and/or controllers, which may take the form of a general or

special-purpose processor or controller. For instance, the processing unit **602** may include microprocessors, micro-controllers, application-specific integrated circuits, digital signal processors, and the like. The memory **604** may be data storage that can be loaded with one or more of the software components executable by the processor **602** to perform those functions. Accordingly, memory **604** may comprise one or more non-transitory computer-readable storage mediums, examples of which may include volatile storage mediums such as random access memory, registers, cache, etc. and non-volatile storage mediums such as read-only memory, a hard-disk drive, a solid-state drive, flash memory, and/or an optical-storage device, among other possibilities.

The microphone array **606** may be a plurality of microphones arranged to detect sound in the environment of the network microphone device **600**. Microphone array **606** may include any type of microphone now known or later developed such as a condenser microphone, electret condenser microphone, or a dynamic microphone, among other possibilities. In one example, the microphone array may be arranged to detect audio from one or more directions relative to the network microphone device. The microphone array **606** may be sensitive to a portion of a frequency range. In one example, a first subset of the microphone array **606** may be sensitive to a first frequency range, while a second subset of the microphone array may be sensitive to a second frequency range. The microphone array **606** may further be arranged to capture location information of an audio source (e.g., voice, audible sound) and/or to assist in filtering background noise. Notably, in some embodiments the microphone array may consist of only a single microphone, rather than a plurality of microphones.

The network interface **608** may be configured to facilitate wireless and/or wired communication between various network devices, such as, in reference to FIG. 5, CR **522**, PBDs **532-538**, computing device **504-508** in cloud network **502**, and other network microphone devices, among other possibilities. As such, network interface **608** may take any suitable form for carrying out these functions, examples of which may include an Ethernet interface, a serial bus interface (e.g., FireWire, USB 2.0, etc.), a chipset and antenna adapted to facilitate wireless communication, and/or any other interface that provides for wired and/or wireless communication. In one example, the network interface **608** may be based on an industry standard (e.g., infrared, radio, wired standards including IEEE 802.3, wireless standards including IEEE 802.11a, 802.11b, 802.11g, 802.11n, 802.11ac, 802.15, 4G mobile communication standard, and so on).

The user interface **610** of the network microphone device **600** may be configured to facilitate user interactions with the network microphone device. In one example, the user interface **608** may include one or more of physical buttons, graphical interfaces provided on touch sensitive screen(s) and/or surface(s), among other possibilities, for a user to directly provide input to the network microphone device **600**. The user interface **610** may further include one or more of lights and the speaker(s) **614** to provide visual and/or audio feedback to a user. In one example, the network microphone device **600** may further be configured to playback audio content via the speaker(s) **614**.

### III. Example Methods

As discussed above, embodiments described herein may involve identifying and accessing suitable streaming services (e.g. streaming audio tracks) based on commands.

Methods **700** and **800** shown in FIGS. 7 and 8 present embodiments of methods that can be implemented within an operating environment involving, for example, the media

playback system **100** of FIG. 1, one or more of the playback device **200** of FIG. 2, and one or more of the control device **300** of FIG. 3. Methods **700** and **800** may include one or more operations, functions, or actions as illustrated by one or more of blocks **702-714** and **802-808**. Although the blocks are illustrated in sequential order, these blocks may also be performed in parallel, and/or in a different order than those described herein. Also, the various blocks may be combined into fewer blocks, divided into additional blocks, and/or removed based upon the desired implementation.

In addition, for the methods **700**, **800**, and other processes and methods disclosed herein, the flowchart shows functionality and operation of one possible implementation of present embodiments. In this regard, each block may represent a module, a segment, or a portion of program code, which includes one or more instructions executable by a processor for implementing specific logical functions or steps in the process. The program code may be stored on any type of computer readable medium, for example, such as a storage device including a disk or hard drive. The computer readable medium may include non-transitory computer readable medium, for example, such as computer-readable media that stores data for short periods of time like register memory, processor cache and Random Access Memory (RAM). The computer readable medium may also include non-transitory media, such as secondary or persistent long term storage, like read only memory (ROM), optical or magnetic disks, compact-disc read only memory (CD-ROM), for example. The computer readable media may also be any other volatile or non-volatile storage systems. The computer readable medium may be considered a computer readable storage medium, for example, or a tangible storage device. In addition, for the methods **700**, **800** and other processes and methods disclosed herein, each block in FIGS. 5 and 6 may represent circuitry that is wired to perform the specific logical functions in the process.

FIG. 7 is an example flow diagram related to a process for identifying a music service for streaming.

At **702**, NMDs **512-516** or CR **522** may receive an indication of a command which may indicate audio content to be provided for playback from a streaming service. In some instances, a command received by NMDs **512-516** may take the form of a voice command, whereas a command received by CR **522** may be a textual command input on a user interface.

Typically, the received command may include information relating to one or more audio content types. In some cases, the command may include the name of an artist, song, album, or genre (i.e. "play Led Zeppelin," "play 70's rock".) Additionally, the command may include pre-fix and/or suffix type information (e.g. "best of . . .," ". . . radio," ". . . playlist") that may be further indicative of content type. For example, a command "play Led Zeppelin radio" may indicate a user's desire to listen to a specific artist's music in a radio format. The received command may include various other forms of information indicative of content type as well.

The received indication of a command at **702** may be processed in various ways. In one implementation, the processing of a command may be accomplished via cloud network **502**. In such a case, a voice command received by NMDs **512-516** may cause the voice input to be transmitted via communication network **546** to one or more of computing device **504-508** for processing. The cloud computing device may convert the voice input to an equivalent text command and parse the text command to identify the command. In another configuration the cloud computing device may only convert the voice input to an equivalent text

format and send the equivalent text to a second computing device for parsing and command identification. In other instances, the NMDs **512-516** may convert the voice input to text prior to transmission via communication network **546** or both convert a voice input to text and perform the parsing to identify the command. In the case of CR **522** receiving a textual command, the text input may be transmitted via communication network **546** to one of computing devices **504-508** for parsing and command identification. In another instance, CR **522** may perform the parsing of the text input to identify the command.

In another implementation, the processing of a command may be accomplished locally over a local network. In such a case, a voice command received by NMDs **512-516** may cause the voice input to be transmitted via a local network to one or more local computing devices for processing. The local computing device may convert the voice input to an equivalent text command and parse the text command to identify the command. In another configuration local computing device may only convert the voice input to equivalent text format and send the the equivalent text format to a second local computing device for parsing and command identification. In other instances, the NMDs **512-16** may convert the voice input to text prior to transmission via the local network or both convert a voice input to text and perform the parsing to identify the command. In the case of CR **522** receiving a textual command, the text input may be transmitted via the local network to a local computing device for parsing and command identification. In another instance, CR **522** may perform the parsing of the text input to identify the command. Other configurations for processing a command may exist.

At **704**, a computing device may identify a content type indicated by the command. The identification of content type may be accomplished via cloud network **502** or locally over a local network.

In one implementation a computing device may use content type logic to correlate the commands or portions of the commands to content type(s). Using the aforementioned example of "Play Led Zeppelin radio," a computing device may identify the content type as "Artist/Radio Station." In another example, the command "Play Electronic Dance Music" may cause the content type to be identified as "Genre." Similar identifications may be made for the various other content types. The identification of content type may be accomplished for instance through inputting a keyword of the command such as "Dance Music" and the database may map the keyword to an indication for content type such as the Genre. The database may reside on the computing device or on the network microphone device in some examples.

If it is determined that no content type is identified at **708** the method may proceed directly to **714**. A content type may be unidentifiable for a number of reasons including user input error, poor speech input quality, background noise, or simply no such content type is known. For example, the content type of a command indicating an obscure artist name may be unable to be identified.

At **714**, an indication may be output by a computing device and transmitted via the communication network **546** to any or all of NMDs **512-516**, PBDs **532-538**, or CR **522** indicating that "the content is unavailable". The indication that no content is available may then be presented audibly or visually to a user. For example, the NMDs and PBDs may output audible indications, whereas the CR may be capable of outputting both audible and visual indications. The indication sent may, additionally or alternatively, cause a suggestion to be output to a user instructing he or she to re-input

the command. For instance, the suggestion might be for the user to specify some additional identifying characteristic so as to assist in identifying the content type.

However, if it is determined at **706** that a content type has been identified the method may proceed to **708** to identify a streaming service that is able to play the content type identified at **704**. Generally, particular streaming services may vary significantly from other streaming services not only in what audio content they provide but also in how they present the content. For instance, each streaming service may possess relatively exclusive rights to stream the music content of certain artists or albums. In another instance, some streaming services, such as Pandora®, may only stream in radio station format, whereas others like Spotify® may be capable of streaming music on demand by artist, song, album, or radio station. In view of this fact, it is apparent that not all streaming services may be capable of streaming a content type identified at **704**.

In one instance, a computing device may identify a suitable streaming service by comparing metadata of the identified content type to a look-up table(s) that may contain entries for the content available and in what format the content is capable of being provisioned for various streaming services. In some cases, the computing device may direct the query to the entire universe of streaming services available. In other cases, the computing device may only query a sub-set of available streaming services. Such a sub-set may be chosen by the computing device based on a number of factors alone or in combination including streaming services a user is registered with, the amount of days since a user has last used a streaming service, streaming service popularity, user settings, among others. For example, if a user has only registered with Pandora®, Spotify®, and Deezer®, the computing device may only query those streaming services to determine which are suitable.

Such look-up table(s) may be stored in memory on a computing device or at an external location such as the computing device or at the music service. Given that the various look-up tables may be distributed amongst a variety of music services, a computing device may query each music service simultaneously or sequentially in order to find a match. Other manners of identifying a suitable streaming service are possible.

In one implementation the identification of a streaming service at **708** may further involve determining a currently available playback capacity of a streaming service that a user is registered with. Generally, some streaming services may limit the number of active streams available for a registered account at any given time. For example, Spotify® may only allow a single active stream per a registered account. In one instance, a computing device may determine the currently available playback capacity by querying the services a user is registered with for a usage status (i.e. how many active streams) and then comparing the usage status to capacity restriction data (i.e. Spotify=1 active stream only). In another example, the streaming services may output a binary value in response to the query to indicate whether or not a stream is available. The available playback capacity may be determined in other ways.

In this implementation, the computing device at **708** may identify a registered service as supporting the content type indicated by the command and further determine the registered service does not have a stream available. For example, if a user and their spouse both share a Spotify® account and music is being streamed to the spouse's smartphone device at the gym when the user issues the command "Play Eye of the Tiger," the computing device may identify Spotify® as

being able to play the song and also that a stream is unavailable. Such a case may cause the computing device to identify another streaming service capable of supporting the content type, such as Apple Music®

In another instance, the computing device may be unable to identify another streaming service that may support the content type. This may occur for example, if a user requests content exclusively provided by a single streaming service, the computing device only considers sub-set of streaming services, among other examples. In such an instance, the computing device may cause a currently active stream to be “stolen” for use in providing the content corresponding to the command. Using the aforementioned, example if no other streaming service is capable of supporting “Eye of the Tiger,” the stream to spouse at the gym may be cancelled and provided to the user.

In one instance, on the occurrence that an available streaming service is identified at 710 as capable of supporting the identified content type, the process may proceed to 714 to cause any combination of PBDs 532-538 to playback the audio content. The music service may be accessed, in one instance through querying the service API for content and causing the content to be streamed. The audio content may be streamed directly from computing device 508 or from various other computing devices associated with streaming music services directly to PBDs 532-538 upon a request from either PBDs 532-538 or computing devices 504-506. Other ways of initiating and causing the playback of streaming media content also exist.

In another instance, if a streaming service identified at 710 is not presently available (i.e. application not installed, user not registered) a computing device may cause, at 714, the output of an indication pertaining to a suggestive course of action to enable the music service to be used. The indication may be sent any combination of NMDs 512-516, PBDs 532-538, or CR 522 and may cause an audible and/or visual suggestion indicating the identified music service capable of supporting the content type and/or present instructions on how to sign up for, download, or otherwise utilize the music service.

FIG. 8 is another example flow diagram related to an example process for identifying a streaming music service in blocks 710 and 712 of FIG. 6.

At 802 a computing devices may cause a confidence metrics or metrics to be determined for streaming service(s) based at least in part on the content type identified by the command. In general, a confidence metric may be a numerical or percentage value (e.g. 1-100) calculated for a streaming service or services. Such confidence metric(s) may reflect the likelihood that a selection of a particular music service for providing streaming audio will result in providing the user with the content he or she desires. For example, a streaming service assigned a confidence metric of 80 may be more suitable for content provision than a streaming service with a confidence metric of 45, where a higher number indicates a higher confidence level of suitability.

The calculated confidence metrics for streaming services may be based on a number criterion such as content type, playback capacity, usage history, external data, among others. Such criteria may be constituted by various data types and may be retrieved from various sources such as the NMDs, CRs, PBDs, computing devices, music services, and various external sources. The data may be aggregated and stored in a central location such as a database associated with computing devices 504 or 506 or in a distributed fashion.

In one instance, the confidence metric may take into account a streaming service’s suitability to support the content type indicated by a command. Determining suitability of various streaming services to provide a content type may involve mapping metadata relating to an identified content type to a look-up table or querying the tables of the various music services and assigning a value to the number of fields that match. As an example, a command that specifies “play Jackson 5 playlist” may have the content type “Artist/playlist.” In such a case, the look-up table(s) of two streaming services such as Pandora® and Spotify® may both contain references to Jackson 5 in an artist field. However, only Spotify® may contain a playlist field identifying Jackson 5, as Pandora® does not support the content type playlist. In such a case, Pandora® may be afforded content type value of 2 and Spotify® a value of 1.

Additionally, or alternatively, strength of field matching may be employed. Using the aforementioned example and assuming that the Spotify® service does not contain an artist field corresponding to Jackson 5 but have artist entry for Michael Jackson, who may be identified by a computing device utilizing music metadata as having been a former member of Jackson 5. In such, a case the Spotify® service may not be given a value of 0 for the artist field, but rather the service may be afforded an adjusted value less than 1. Other forms of determining streaming service content type suitability value are possible.

In another instance, the confidence metrics may be calculated in part on various forms of historical usage data. The various historical data types may be retrieved from various sources such as the NMDs, CRs, PBDs, computing devices, music services, and various external sources. The data may be aggregated and stored in a central location such as a database associated with computing devices 504 or 506 or in a distributed fashion.

For example, the usage data may indicate how frequently a user accesses a given streaming service. As another example, the usage data may include time based data to identify the service a user normally uses at various times of the day, days of the week, and months of the year. For example, a user may prefer to listen to iHeartRadio® in the morning and to Tidal® in the evenings. As another example, in a multi-zone environment such usage data may indicate user preferences regarding streaming services on a per zone basis. For example, if a user typically accesses Spotify® 90% of the time in a bathroom zone and Deezer® 80% of the time in a living area zone the confidence metrics corresponding to Spotify® and Deezer® may vary significantly depending on which zone the user intends to stream audio to. In such, a case if the user desired to listen to music in the bathroom, Spotify® would be the much better choice. Various other types of historical usage data may exist as well.

Additionally, the confidence metrics may be determined in part based on various “external” data types. Such data may include macro type data that may take into account geographic location or demographic data, among other possibilities. For example, such macro data may indicate that a particular streaming service is not available or is unpopular in certain regions of the world or sub-regions of a country, which may result in a lower confidence metric. In such a case, the confidence metric for that streaming service may vary dependent on the geographic location. Types of “external” data may further include weather data, which may be taken into account, for example a user’s preference to listen to Spotify® on the patio on cool, dry summer evenings. Additionally, calendar data may be considered to identify

holidays and the music service typically streamed on those days. Other forms of external data may exist.

Furthermore, it is possible to combine the various criteria to determine a confidence metric for a given streaming service. For instance, a user may prefer to use Pandora® to listen to a wide variety of classical music in the living room but may exclusively use Apple Music® to listen to full albums of their favorite artist in the bedroom. In such a case, content type(s) (genre, artist) may be combined with a user history relating to location of use of particular music services to determine a confidence metric for a particular music service. Numerous other combinations may exist.

A computer implemented algorithm may map the criterion described above to a confidence metric for streaming service. For instance, one or more criterion may be mapped into a table which outputs the confidence metric. In some embodiments, the algorithm may involve weighting of various criteria such as content types, usage history, and/or "external data." The algorithm may assign different weights to the various criteria based on relative importance. For example, a user preference may be deemed more influential and afforded more weight than demographic data. The weighting of inputs may be defined by the system, by user settings, or adjustable dynamically based on user feedback. Each of weighted inputs may be input into a table, for example, for mapping to confidence metrics which are then combined to form an aggregated confidence metric for a music service as a whole.

At **804** either computing device **504** or **506** may determine whether a given streaming service satisfies a confidence condition. The confidence condition may be satisfied in a number of ways. In one instance, the confidence condition may be satisfied if the confidence metric for a given streaming service exceeds a confidence level threshold, which may be a default system setting or adjustable by a user. For example, if three confidence metrics are calculated as Service1=85, Service2=83, Service3=25 with the threshold confidence level being 80, may cause Service1 to be output for streaming or suggested to a user at **806** (as discussed above in reference to **714**).

In another instance, the confidence condition may only be satisfied upon (1) a streaming service with a calculated confidence metric above the threshold confidence level and (2) the two highest calculated confidence metrics are not within a threshold range of one another. For example, if the confidence level range were 3 in the case mentioned directly above the confidence condition would not be satisfied due to confidence metrics of Service1 and Service2. The confidence condition may additionally or alternatively incorporate various other rules.

At **808** an error state may be triggered by either computing device **504** or **506** if the confidence condition is found not to be satisfied at **808**. An error state may trigger certain events to be caused, such as outputting an indication that the content is unavailable, as discussed in reference to **716**. In another instance, a triggered error state may cause content to be streamed via a default streaming service or a preferred partner service. In yet another instance, the error state may cause a computing device to output an indication to one or all of NMDs **512-516**, PBDs **532-538**, or CR **522** to cause an audible or visual presentation of an instruction or query directed at obtaining more information in relation to the already received command.

In one instance, a user may be instructed to provide an additional content type such as artist or album. For example, if an initial command were "Play radio station" a user may be instructed "Please provide genre."

In another instance, a user may be asked a question or series of questions that may help further tune the initial command and thereby the confidence metrics for the various streaming services. For example, if an initial command were directed to the the broad genre of "Electronic," the user may be asked whether they prefer certain sub-genres such as "Drum and Bass" or "Trance." Additionally or alternatively, the user may be asked questioned unrelated to music such as "What are you doing?" or "How do you feel?" to infer what sub-genre a user may want to listen to. For example, if the user answered "Reading" and/or "Relaxed" to the aforementioned questions it may be inferred that the user is interested in the sub-genre Trance. Various other examples of queries are possible.

The user responses to the instructions or questions may take the form of speech input which may be received by an NMD or a textual input via a graphical interface of a CR. The user responses may cause process **700** to repeat which may result in the confidence condition being satisfied or an additional error state being triggered.

#### IV. Conclusion

The description above discloses, among other things, various example systems, methods, apparatus, and articles of manufacture including, among other components, firmware and/or software executed on hardware. It is understood that such examples are merely illustrative and should not be considered as limiting. For example, it is contemplated that any or all of the firmware, hardware, and/or software aspects or components can be embodied exclusively in hardware, exclusively in software, exclusively in firmware, or in any combination of hardware, software, and/or firmware. Accordingly, the examples provided are not the only way(s) to implement such systems, methods, apparatus, and/or articles of manufacture.

Additionally, references herein to "embodiment" means that a particular feature, structure, or characteristic described in connection with the embodiment can be included in at least one example embodiment of an invention. The appearances of this phrase in various places in the specification are not necessarily all referring to the same embodiment, nor are separate or alternative embodiments mutually exclusive of other embodiments. As such, the embodiments described herein, explicitly and implicitly understood by one skilled in the art, can be combined with other embodiments.

The specification is presented largely in terms of illustrative environments, systems, procedures, steps, logic blocks, processing, and other symbolic representations that directly or indirectly resemble the operations of data processing devices coupled to networks. These process descriptions and representations are typically used by those skilled in the art to most effectively convey the substance of their work to others skilled in the art. Numerous specific details are set forth to provide a thorough understanding of the present disclosure. However, it is understood to those skilled in the art that certain embodiments of the present disclosure can be practiced without certain, specific details. In other instances, well known methods, procedures, components, and circuitry have not been described in detail to avoid unnecessarily obscuring aspects of the embodiments. Accordingly, the scope of the present disclosure is defined by the appended claims rather than the forgoing description of embodiments.

When any of the appended claims are read to cover a purely software and/or firmware implementation, at least one of the elements in at least one example is hereby expressly defined to include a tangible, non-transitory medium such as a memory, DVD, CD, Blu-ray, and so on, storing the software and/or firmware.

We claim:

1. A method comprising:

receiving, by a computing device, data indicating a request for audio content by a media playback system, wherein the media playback system comprises a plurality of playback zones, wherein each playback zone comprises at least one playback device, and wherein each playback device is configured to (i) receive and playback audio content in a radio format from a first music service provider and (ii) receive and playback audio content in a non-radio format from a second music service provider or a third music provider;

parsing by the computing device, the data to identify at least (i) a first segment indicating a content identification, (ii) a second segment indicating one of the radio format and the non-radio format, and (iii) a third segment indicating a particular playback zone of the plurality of playback zones;

selecting by the computing device, one of the first music service provider, the second music service provider, and the third music provider based on the second segment and the third segment, wherein if the second segment indicates the radio format, the computing device selects the first music service provider, and wherein if the second segment indicates the non-radio format, the computing device selects one of the second music service provider and the third music provider, wherein selecting one of the second music service provider and the third music service provider comprises:

determining a first confidence metric corresponding to the second music service provider and a second confidence metric corresponding to the third music service provider, wherein the first confidence metric is based on a history of accesses to the second music service provider at the particular playback zone of the plurality of playback zones indicated in the third segment and wherein the second confidence metric is based on a history of accesses to the third music service provider at the particular playback zone of the plurality of playback zones indicated in the third segment and

determining whether the first confidence metric or the second confidence metric satisfies a confidence condition;

after selecting one of the first music service provider, the second music service provider, and the third music provider based on the second segment and the third segment, identifying a uniform resource identifier corresponding to the content identification and the selected music service provider;

identifying a particular playback device in the particular playback zone indicated in the third segment; and

transmitting by the computing device to the particular playback device, the uniform resource identifier to cause the at least one playback device to retrieve audio content from the selected music service provider according to the uniform resource identifier.

2. The method of claim 1, wherein the data indicating a request for audio content by a media playback system is received via a network microphone device.

3. The method of claim 1, further comprising determining that the second music service provider does not support audio content in the non-radio format.

4. The method of claim 1, wherein the content identification indicates one or more of an Artist, Genre, Song, and Album of the requested audio content.

5. A network device comprising:

a network interface configured to communicate with a plurality of networked devices over a network;

a processor comprising instructions, which when executed, cause the processor to:

receive via the network interface, data indicating a request for audio content by a media playback system, wherein the media playback system comprises a plurality of playback zones, wherein each playback zone comprises at least one playback device, and wherein each playback device is configured to (i) receive and playback audio content in a radio format from a first music service provider and (ii) receive and playback audio content in a non-radio format from a second music service provider or a third music provider;

parse the data to identify at least (i) a first segment indicating a content identification, (ii) a second segment indicating one of the radio format and the non-radio format, and (iii) a third segment indicating a particular playback zone of the plurality of playback zones;

select one of the first music service provider, the second music service provider, and the third music provider based on the second segment and the third segment, wherein if the second segment indicates the radio format, the first music service provider is selected, and wherein if the second segment indicates the non-radio format, one of the second music service provider and the third music provider is selected, wherein selecting one of the second music service provider and the third music service provider comprises:

determine a first confidence metric corresponding to the second music service provider and a second confidence metric corresponding to the third music service provider, wherein the first confidence metric is based on a history of accesses to the second music service provider at the particular playback zone of the plurality of playback zones indicated in the third segment and wherein the second confidence metric is based on a history of accesses to the third music service provider at the particular playback zone of the plurality of playback zones indicated in the third segment and

determine whether the first confidence metric or the second confidence metric satisfies a confidence condition;

after selection of one of the first music service provider, the second music service provider, and the third music provider based on the second segment, identify a uniform resource identifier corresponding to the content identification and the selected music service provider;

identify a particular playback device in the particular playback zone indicated in the third segment; and transmit via the network interface to the particular playback device, the uniform resource identifier to cause the at least one playback device to retrieve audio content from the selected music service provider according to the uniform resource identifier.

6. The network device of claim 5, wherein the data indicating a request for audio content by a media playback system is received via a network microphone device.



25

7. The network device of claim 5, further comprising instructions to determine that the second music service provider does not support audio content in the non-radio format.

8. The network device of claim 5, wherein the content identification indicates one or more of an of Artist, Genre, Song, and Album of the requested audio content.

9. A tangible, non-transitory computer readable storage medium including instructions for execution by a processor, the instructions, when executed cause the processor to implement a method comprising:

receiving data indicating a request for audio content by a media playback system, wherein the media playback system comprises a plurality of playback zones, wherein each playback zone comprises at least one playback device, and wherein each playback device is configured to (i) receive and playback audio content in a radio format from a first music service provider and (ii) receive and playback audio content in a non-radio format from a second music service provider or a third music provider;

parsing the data to identify at least (i) a first segment indicating a content identification, (ii) a second segment indicating one of the radio format and the non-radio format, and (iii) a third segment indicating a particular playback zone of the plurality of playback zones;

selecting one of the first music service provider, the second music service provider, and the third music provider based on the second segment, wherein if the second segment indicates the radio format, the first music service provider is selected, and wherein if the second segment indicates the non-radio format, one of the second music service provider and the third music provider is selected, wherein selecting one of the second music service provider and the third music service provider comprises:

26

determining a first confidence metric corresponding to the second music service provider and a second confidence metric corresponding to the third music service provider, wherein the first confidence metric is based on a history of accesses to the second music service provider at the particular playback zone of the plurality of playback zones indicated in the third segment and wherein the second confidence metric is based on a history of accesses to the third music service provider at the particular playback zone of the plurality of playback zones indicated in the third segment and

determining whether the first confidence metric or the second confidence metric satisfies a confidence condition;

after selecting one of the first music service provider, the second music service provider, and the third music provider based on the second segment, identifying a uniform resource identifier corresponding to the content identification and the selected music service provider;

identifying a particular playback device in the particular playback zone indicated in the third segment; and transmitting to the particular playback device, the uniform resource identifier to cause the at least one playback device to retrieve audio content from the selected music service provider according to the uniform resource identifier.

10. The tangible, non-transitory computer readable storage medium of claim 9, wherein the data indicating a request for audio content by a media playback system is received via a network microphone device.

11. The tangible, non-transitory computer readable storage medium of claim 9, wherein the method further comprises determining that the second music service provider does not support audio content in the non-radio format.

\* \* \* \* \*